Metabarcoding and metabolome analysis of copepod grazing reveals feeding preference and linkage to metabolite classes in dynamic microbial plankton communities

Jessica L. Ray[1*], Julia Althammer[2□], Katrine S. Skaar[1], Paolo Simonelli[3], Aud Larsen[1], Diane Stoecker[4], Andrey Sazhin[5], Umer Z. Ijaz[6], Christopher Quince[7], Jens C. Nejstgaard[8], Marc Frischer[1,9], Georg Pohnert[2] and Christofer Troedsson[1]

[1]Hjort Centre for Marine Ecosystem Dynamics, Uni Research Environment, Uni Research AS, Nygårdsgaten 112, N-5008 Bergen, Norway

[2]Institute for Inorganic and Analytical Chemistry, Friedrich Schiller University Jena, Lessingstr. 8, 07443 Jena, Germany

[3]Department of Biology, University of Bergen, Thormøhlensgt 53A, 5006 Bergen, Norway

[4]University of Maryland, Center of Environmental Science, Horn Point Lab, Cambridge, Massachusetts, 21613, USA

[5]P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences, Laboratory of Ecology of Plankton Organisms, Nakhimovsky Prospect 36, Moscow, Russia

[6]School of Engineering, University of Glasgow, Glasgow G12 8QQ, United Kingdom

[7]WMS - Microbiology and Infection, University of Warwick Medical School, Coventry, CV4 7AL, United Kingdom

[8]Leibniz Institute of Freshwater Ecology and Inland Fisheries, Dep. 3, Experimental Limnology/Alte Fischerhütte 2, 16775 Stechlin, OT Neuglobsow, Germany

[9]Skidaway Institute of Oceanography, 10 Science Circle, Savannah, Georgia 31411, USA

Keywords: *Calanus* grazing, *Phaeocystis pouchetii*, *Skeletonema marinoi*, seawater mesocosms, eukaryote V7 SSU, metabolome profiling, metabarcoding, symbiosis

Running Head: Multi-omic analysis of *Calanus* feeding

[*]Corresponding author: E-mail jessicalouiseray@gmail.com, Telephone (+47) 55 58 26 73, Fax (+47) 55 58 96 74

□ Current address: Jenabios GmbH, Orlaweg 2, 07743 Jena, Germany.

Abstract

In order to characterize copepod feeding in relation to microbial plankton community dynamics, we combined metabarcoding and metabolome analyses during a 22-day seawater mesocosm experiment. Nutrient amendment of mesocosms promoted the development of haptophyte-

(*Phaeocystis pouchetii*) and diatom- (*Skeletonema marinoi*) dominated plankton communities in mesocosms, in which *Calanus* sp. copepods were incubated for 24-hours in flow-through chambers to allow access to prey particles (< 500 µm). Copepods and mesocosm water sampled six times spanning the experiment were analyzed using metabarcoding, while intracellular metabolite profiles of mesocosm plankton communities were generated for all experimental days. Taxon-specific metabarcoding ratios (ratio of consumed prey to available prey in the surrounding seawater) revealed diverse and dynamic copepod feeding selection, with positive selection on large diatoms, heterotrophic nanoflagellates and fungi, while smaller phytoplankton, including *P. pouchetii*, were passively consumed or even negatively selected according to our indicator. Our analysis of the relationship between *Calanus* grazing ratios and intracellular metabolite profiles indicates the importance of carbohydrates and lipids in plankton succession and copepod-prey interactions. This molecular characterization of *Calanus* sp. grazing therefore provides new evidence for selective feeding in mixed plankton assemblages and corroborates previous findings that copepod grazing may be coupled to the developmental and metabolic stage of the entire prey community rather than to individual prey abundances.

Introduction

The trophic efficiency of the marine food web depends upon the pathway of carbon flow from primary production to predatory fish - either through the classical food web (diatoms to mesozooplankton), the microbial food web (flagellates/bacteria to ciliates to mesozooplankton) (Landry 2002; Pepin *et al*. 2011), or via the recently discussed nutritunneling to bypass phytoplankton pathways (Pitta *et al.* 2016). Copepods are among the most abundant mesozooplankton in the global ocean and have long been assumed to be key regulators of carbon

transfer through their selective feeding in the marine food web (Sherr & Sherr 1988; Kleppel 1993; Sanders & Wickham 1993). Understanding the interplay of copepod grazing as a top-down regulatory force, in concert with the bottom-up regulation of marine plankton assemblages, is essential for accurately predicting the flow of carbon and nutrients from primary production to fisheries (Sherr & Sherr 1988; Calbet & Saiz 2005).

Copepods can be selective grazers (Kiørboe *et al.* 1996; Kiørboe 2011 and references therein) dependent upon prey abundance, size, motility or chemical cues (Nejstgaard *et al.* 2008 and references therein). Elucidating copepod prey selection in natural environments presents a persistent methodological challenge, as full characterization of copepod feeding requires quantitative knowledge of the potential prey community as well as knowledge of the prey organisms that are actually consumed (reviewed in Pompanon *et al*. 2011). Chlorophyll *a* or pigment measurements of copepod gut content cannot be used to determine prey selection in natural prey assemblages as pigments show variable breakdown rates and do not reveal non-pigmented prey that may frequently be the most selected prey *in situ* (see data for *Calanus* in Nejstgaard *et al.* 1997; 2001b, and further discussion on methods in Nejstgaard *et al*. 2008). Molecular analysis of phylogenetic markers (metabarcoding; Taberlet *et al.* 2012) provides a promising alternative due to the universality of genomic DNA among cellular organisms and tunable precision of phylogenetic resolution without previous knowledge of community composition. In addition, tools exist to refine the output of molecular investigation toward a more prey-oriented analysis through the use of DHPLC-PCR (Troedsson *et al*. 2008a,b; Olsen *et al*. 2014), restriction endonucleases (Maloy *et al.* 2013) and/or blocking oligonucleotides (Troedsson *et al.* 2008b; Vestheim & Jarman 2008; Maloy *et al.* 2013) to selectively inhibit

amplification of predator sequences.

Ample evidence exists to suggest that chemical cues likely play an equitable role with prey availability in copepod feeding behavior (Poulet & Marsot 1978; Cowles *et al.* 1988). The combination of DNA-based metabarcoding methods with intracellular metabolite profiling may thus facilitate a deeper investigation of copepod feeding behavior (Woodson *et al.* 2007) that takes into account both taxonomy and chemical ecology (Barofsky *et al.* 2010; Kuhlisch & Pohnert 2015). We employed a multi-omic approach to characterize microbial plankton communities in seawater mesocosms dominated by the haptophyte *Phaeocystis pouchetii* or the diatom *Skeletonema marinoi*, and compared seawater microplankton communities to the "community" of prey organisms in the gut content of *Calanus* sp. copepods. The aim of this study was to utilize molecular proxies for feeding selection generated by metabarcoding analysis to investigate whether changes in *Calanus* feeding selection could be linked to changes in the chemical profile of co-occurring microbial planktonic communities.

Materials and Methods

Mesocosm experiment

A seawater mesocosm experiment was performed during a 22-day period from 8-30 March 2012 at the Espegrend Marine Biological Station at the University of Bergen, Norway. A detailed description of the experimental set-up is provided elsewhere (Nejstgaard *et al.* 2006; Stoecker *et al.* 2015; Ray *et al.* 2016). Briefly, triplicate 11 m^3 reinforced transparent polyurethane mesocosms bags were either left unamended (Control), amended with 16 µM NO_3^- and 1 µM PO_4^{3-} (NP) to selectively promote *P. pouchetii* growth, or amended with 16 µM NO_3^-, 1 µM

105 PO_4^{3-} and 5 µM SiO_4^{2-} (NPSi) to selectively promote diatom growth. For reference,
106 unmanipulated samples were also taken from Raunefjorden (Raunefjorden) directly adjacent to
107 the mesocosm raft. A detailed description of mesocosm bloom development is available
108 elsewhere (Stoecker *et al.* 2015; Ray *et al*. 2016). Briefly, exponential growth of the diatom *S.*
109 *marinoi* occured in all three mesocosm treatments (Control, NP and NPSi), peaking during 18-23
110 March. Highest abundances of *S. marinoi* occurred in the NPSi treatment. Exponential growth of
111 *P. pouchetii* blooms commenced after 25 March in both the NP and NPSi mesocosms, although
112 highest abundances of *P. pouchetii* occurred in the NP mesocosms (Ray *et al.* 2016). The
113 initially high nutrient levels in the Raunefjorden, from which mesocosm bags were filled,
114 resulted in similar successive blooms in the different mesocosm treatments despite differential
115 nutrient amendment (Stoecker *et al*. 2015).
116
117 **Sampling for metabarcoding analysis**
118 Samples for metabarcoding were collected on 11, 17, 21, 24, 28 and 30 March 2012 as described
119 previously (Ray *et al.* 2016, Table 1). Briefly, triplicate 50-200 ml seawater samples from all
120 mesocosms and Raunefjorden were filtered by gentle vacuum onto 0.2 µm SUPOR filters (Pall
121 Corporation). Filters were aseptically transferred to 2.0 ml tubes containing 280 µl of 56ºC
122 Buffer ATL and 20 µl Proteinase K (20 mg ml^{-1}) (QIAGEN DNeasy Blood & Tissue kit). For
123 feeding chamber incubations, *Calanus* sp. were first collected by net tows from Raunefjorden
124 (60°16'18"N, 5°10'26"E), and 20 individuals each of adult female or stage V copepodites were
125 manually sorted into ~ 1.8 L volume grazing chambers containing 0.2 µm-filtered seawater. The
126 detailed chamber construction is described in Ray *et al.* (2016). Sorted copepods inside grazing
127 chambers were kept in the dark at 8°C until deployment inside mesocosms on the following
128 morning. Three replicate grazing chambers containing copepods were deployed inside each

mesocosm. The 500 µm nylon mesh covering both openings of the grazing chambers allowed constant vertical circulation of mesocosm water containing *in situ* microbial assemblages through the chambers. After a 24-hour incubation, mesocosm chambers were recovered one at a time, and copepods were rinsed with three successive washes of 0.2 µm-filtered seawater prior to a final immersion in an anaesthetic seawater solution containing 0.37 mg ml^{-1} ethyl 3-aminobenzoate methanesulfonate (MS222) (Sigma-Aldrich, Norway) (Simonelli *et al.* 2009). Pools of five copepods from each chamber were sorted into 1.5 ml tubes containing 180 µl of Buffer ATL preheated to 56ºC and 20 µl (20 mg ml^{-1}) Proteinase K (QIAGEN DNeasy Blood & Tissue kit). In summary, we collected three biological replicate copepod samples (five copepod individuals per sample) and three biological replicate seawater samples (filters) per mesocosm and from Raunefjorden on each of the six sampling days. Because whole copepods were lysed, the copepod samples may also contain any symbionts on copepod surfaces or in tissues. For simplicity, however, we refer to these samples as "copepod gut content" samples throughout the manuscript. All filter and copepod samples for DNA extraction were lysed at 56ºC overnight then frozen at -20ºC until processing. DNA extraction was performed according to the manufacturer's protocol, except that two rounds of elution with 100 µl of 56ºC Buffer AE (QIAGEN) were used instead of the recommended single elution step using 200 µl of room temperature Buffer AE.

Amplicon library preparation

The universal primers F-1183mod and R-1443mod (Table 1) targeting the V7 region of the small subunit ribosomal RNA (SSU rRNA) gene were used in polymerase chain reactions (PCR) to amplify microbial eukaryotes. Universality of primers was checked using the TestPrime function on the Arb-Silva website (http://www.arb-silva.de/search/testprime/) (Table S1). Primer F-

1183mod was modified with the Roche GS-FLX Lib-L Adapter B sequence (5' - CCT ATC CCC TGT GTG CCT TGG CAG TC - TCAG - primer - 3') and used in all PCR reactions. Unidirectional sequencing was performed from primer R-1443mod, which was modified with Adapter A (5' - CCA TCT CAT CCC TGC GTG TCT CCG AC -TCAG - barcode - primer - 3'), where a unique 10-mer barcode multiplex identifier (MID) was included for sample identification during demultiplexing. Primers were HPLC-purified to ensure uniform length and to eliminate contaminating DNA from lyophilized primer preparations. In order to block *Calanus* amplification from copepod samples, a combination of the *Calanus*-specific blocking oligos, Cal-SpcC3-block and Cal-PNA-block (Table 1 and Supporting Information) was used in 50 µl PCR reactions containing 5 µl of template DNA, 1X HF buffer (New England Biolabs, Ipswich, Massachusetts), 0.4 U Phusion DNA polymerase (New England Biolabs), 50 µM each dNTP, 250 nM each primer, 2 µM Cal-SpcC3-block and 1 µM Cal-PNA-block. For amplification from seawater samples, no blocking oligos were included but otherwise PCR conditions were identical. Amplification was performed using a C-1000 thermal cycler (Bio-Rad) with one cycle of 95ºC for 5 min, 30 cycles of 94ºC for 20 sec/70.2ºC for 10 sec/60.2ºC for 20 sec/72ºC for 25 sec, and a final elongation of 72ºC for 2 min. PCR products were pooled by sample, purified using 0.8 vol magnetic beads (Agencourt Ampure XP, Beckman-Coulter Inc., Indianapolis, Indiana, USA), and quantified using PicoGreen (Quant-It PicoGreen dsDNA Assay kit, Life Technologies Ltd, Paisley, UK) on a NanoDrop 3300 fluorospectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts). Equimolar amounts of each sample pool were combined and vacuum-concentrated to generate one amplicon library from copepod samples and one from seawater samples. A detailed description of amplicon library generation for metabarcoding may be found in the Supporting Information. Amplicon libraries were sequenced on ½ plate each

using Roche GS-FLX Titanium chemistry at the Norwegian Sequencing Centre (University of Oslo, Norway).

Sequence analysis

We used two different approaches for analysis of metabarcoding results - operational taxonomic unit (OTU) clustering and taxonomic classification. In order to identify taxa and OTUs common to both libraries, sequence data for the two data sets (copepod gut content and seawater amplicon libraries) were collated prior to taxonomic classification or to OTU clustering using 98% sequence similarity cut-off. OTU clustering with 98% cut-off was performed using AmpliconNoise v.1.29 (Quince *et al.* 2009; 2011). Alternatively, forward primer trimming and quality filtering using trim.seqs in mothur v.1.33 (Schloss *et al.* 2009) was performed with the following parameters: qaverage=25, maxambig=0, maxhomop=6, minlength=200, flip=T, pdiffs=1 with raw sequence fasta file and qfile as input. Taxonomic classification of these quality-filtered reads was performed using the CREST classifier with the SilvaMod database as reference set (Lanzén *et al.* 2012). CREST taxonomic classification results for individual samples files were collated using custom awk scripts. Because the SilvaMod database uses the Silva taxonomy without a strict taxonomic ranking, and because the CREST classifier uses a lowest common ancestor (LCA) algorithm, we refer to taxonomic assignments according to rank rather than according to standard taxonomic hierarchy.

Metabolite profile analysis

Phytoplankton samples for intracellular metabolic profiling were collected every day from 8 to 30 March from all mesocosms and from Raunefjorden and analyzed using a modified protocol for metabolomic analysis of *Skeletonema marinoi* cultures (Vidoudez & Pohnert 2012). Briefly,

cells from 9 L of seawater were concentrated onto GF/C-filters (1.2 µm nominal pore size) with gentle vacuum (~400 mbar). Depending on chlorophyll *a* (chl*a*) concentration and microscopy counts, the daily sampling volume was gradually decreased to 1 L during the experiment. Samples were extracted, derivatized and analysed as described previously (Vidoudez & Pohnert 2012). A detailed description of methods used for metabolomic profiling is provided in the Supporting Information.

Statistical analysis

All statistical analyses, unless otherwise noted, were conducted in the R statistical computing environment (R Core Team, 2015). Alpha- and beta-diversity estimates were performed using the rarecurve and estimateR (richness) and diversity functions in the "vegan" package v.2-0.10 (Oksanen *et al.* 2013). Non-metric multidimensional scaling (NDMS) ordination of OTU and taxonomy tables was performed using the metaMDS function in "vegan". Data visualization was performed using the "ggplot2" package (Wickham 2009). Permutational analysis of variance using distance matrices (PERMANOVA) was performed using the adonis function in "vegan". Constrained correspondence analysis (CCA) was performed using the cca function in "vegan" with ordistep for forward selection to identify significant environmental variables. Correlation analysis of metabarcoding ratios to individual metabolites was performed using cor.test (method="kendall") and custom R scripts. The data matrix from metabolomic runs was analysed using canonical analysis of principle coordinates by linear discriminant analysis (CAPdiscrim; Anderson & Willis 2003). CAPdiscrim was performed using the Windows-executable program CAP12 (Anderson 2004) using the following parameters: choice of transformation=none, choice of standardisation=none, choice of distance measure=Bray-Curtis, type of analysis=discriminant analysis, number of principal coordinate axes chosen by the program, number of random

permutations=999. Treatment of metabolomics runs to achieve the data matrix is described in the Supporting Information.

Results

Metabarcoding analysis

Unidirectional pyrosequencing of the seawater amplicon libraries resulted in 437 522 sequence reads, while the copepod gut content libraries generated 353 459 sequence reads (Table S2). Read coverage for each sample was relatively even for the seawater library, with a maximum approximate two-fold variation in number of reads per sample from highest to lowest coverage and a relatively low standard deviation for both pipelines (Table S2). The *Calanus* gut content sample coverage was less even, with a nearly nine-fold difference in the number of sequence reads per sample between samples with lowest and highest sequence coverage (Table S2). Denoising, quality filtering and OTU clustering and chimera removal using AmpliconNoise generated 3115 OTUs in total for both amplicon libraries, while taxonomic assignment of denoised, quality-filtered reads using the CREST classifier and the SilvaMod database as reference resulted in the identification of 1032 unique taxonomic rankings (herein referred to as taxa) in the collated seawater and *Calanus* gut content sequence data (Table S2). Initial inspection of OTU clustering and taxonomy results for the sequence data showed that many reads (2.9 - 36.6 %) present in the *Calanus* gut content amplicon libraries had highest similarity to *Calanus* copepod SSU sequences (black bars in Fig. S1). *Calanus*-like sequences were assumed to be copepod sequences that were amplified despite the blocking-PCR strategy used, and were therefore removed from the copepod gut content OTU and taxonomy tables prior to diversity or statistical analysis (Table S2, footnote 2). It should be noted that amplification strategy used is unable to distinguish between (1) *Calanus* sp. sequences that are the result of

incomplete blocking of predator DNA amplification and (2) *Calanus* sp. sequences that originate from ingestion of other calanoid species, nauplii and/or eggs. Assuming the former and thus removing these sequences prior to downstream analysis, we have therefore not assessed the contribution of cannibalism or predation on closely-related *Calanus* taxa.

Rarefaction analysis of AmpliconNoise OTUs (Fig. S2A, B) and CREST taxonomic assignments (Fig. S2C, D) demonstrated clear undersampling for all *Calanus* gut content samples (Fig. S2A, C), and in particular for those samples for which relatively few reads were obtained (Table S3). Seawater samples, however, were sampled to near saturation (Figures S2B, D), with the exception of seawater samples from Raunefjorden on 17 and 21 March 2012. Closer inspection of sequence data from these two seawater samples showed a larger number of singletons (data not shown) and higher biodiversity (Table S3) relative to the other seawater samples. Beta diversity analysis of abundance-normalized OTU and taxonomy tables based on the inverse Simpson's diversity index (1/D) for Jaccard ($1/D_J$) and Bray-Curtis ($1/D_{BC}$) distance matrices suggested low variation in diversity between seawater samples regardless of analysis pipeline or distance metric used (Table S3, Mesocosm seawater). For *Calanus* gut content samples (from which *Calanus*-like reads were removed), however, we observed higher variation in beta diversity when Bray-Curtis was used as distance metric, but not when Jaccard distances were used (Table S3, *Calanus* gut content). In addition, we observed a trend of increasing beta diversity toward the end of the mesocosm experiment for *Calanus* gut content samples for all treatments, regardless of analysis pipeline (Table S3, *Calanus* gut content).

The diversity revealed by metabarcoding of seawater microbial plankton includes all major marine plankton groups, including rhizarians, diatoms, haptophytes, ciliates, dinoflagellates,

chlorophytes, cercozoans, choanoflagellates, cryptophytes, fungi, bivalves, gastropods, tunicates, crustaceans, land plants and others, indicating that the PCR primers used amplify DNA from a broad diversity of eukaryotic microorganisms present in the Raunefjorden ecosystem (Fig. 1). The taxonomic diversity observed was highly uneven, with the 20 most abundant OTUs (Fig. 1A) or taxa (fifth-rank taxonomic assignments according to the Silva taxonomy) (Fig. 1B) comprising 20-95% of reads in both seawater and copepod gut content sequence libraries. Inspection of taxonomic assignments for the most abundant OTUs/taxa observed in copepod gut content revealed that these reads may represent organisms living in a symbiotic relationship (*sensu* Leung and Polin 2008) with *Calanus*, including Syndiniales (Dinophyceae), Oligohymenophorea (Ciliophora), *Amoebophyra* (Ciliophora), *Paradinium* (Rhizaria). Lacking proof, however, that these reads represent symbioses rather than ingested prey, and because these organisms would likely generate a chemical signal that would be detected by the metabolite profiling analysis, we chose to retain them in the sequence data for downstream statistical analyses.

Non-metric multidimensional scaling (NMDS) analysis of OTU or taxonomic classification results demonstrated that the strongest effect on sample diversity was sample type, that is, whether the sample originated from seawater or from *Calanus* gut content (Fig. 2). This significance was confirmed by PERMANOVA analysis of OTU (AmpliconNoise) and taxonomic (CREST) diversity, for which Pr (>F) = 0.001 for both pipelines independent of distance metric (Table 2). Indeed, sample type explained 21-32.1% of the variation in OTU diversity or 34-47% of the variation in the taxonomic diversity (Table 2) observed. PERMANOVA also identified sampling date as a significant explanatory variable for both AmpliconNoise OTU (6.8 - 8.8%) and CREST taxonomic (7.6 - 9.6%) diversity (Table 2).

Experiment type (mesocosm or fjord) explained a significant fraction of diversity for AmpliconNoise OTUs (3.8 - 4.6% of OTU diversity) but not for CREST taxonomic diversity (< 1.5%). Treatment was not found to be significant regardless of pipeline or distance metric (Table 2). Kruskal-Wallis rank sum tests identified several taxa (Fig. S3) that were significantly different in their relative abundance between seawater and *Calanus* gut content samples. In general, copepod samples were distinguishable by their significantly higher abundances of protozoans, primarily apostome ciliates and the rhizarian *Paradinium*. For two representative taxa, *Phaeocystis pouchetii* and *Skeletonema marinoi*, which exhibited exponential growth during the mesocosm experiment, metabarcoding signals were clearly stronger in seawater samples than in the corresponding copepod samples (Fig. 1 and Fig. S3). Seawater samples also contained significantly higher numbers of dinoflagellates, cercozoans, nanophytoplankton (Phaeocystales, Bacillariophyta, Prymnesiales, Pelagophyta, Chlorophyta), heterotrophic nanoflagellates and radiolarians relative to the corresponding copepod samples (Fig. S3). Copepod gut content samples, on the other hand, contained relatively higher numbers of ciliates, fungi, arthropods, and kinetoplastid protozoans (Fig. S3).

Molecular proxy for *Calanus* feeding selection

The quantitative power of the metabarcoding data was examined by comparison with previous results from qPCR and microscopy analysis of *P. pouchetii* and *S. marinoi* performed on samples from the same mesocosm experiment (Ray *et al*. 2016) (Fig. 3). Metabarcoding signal dynamics in seawater were generally consistent with both microscopic enumeration and qPCR quantification of these taxa in each treatment, although the decreasing microscopy and qPCR signals for *S. marinoi* at the end of the experiment in the NP and NPSi treatments were not evident in the metabarcoding data (Fig. 3B). In concert, however, these results support a semi-

quantitative interpretation of metabarcoding data to evaluate the dynamics of *Calanus* grazing responses to individual prey taxa as a function of specific prey abundance (D'Amore *et al.* 2016; Lanzén *et al.* 2016). The ratios of individual OTUs or taxa abundance in copepod gut content to their abundance in surrounding seawater were therefore calculated as a proxy, or selectivity index (Irigoien *et al.* 2000), for *Calanus* grazing (Ray *et al*. 2016) on individual OTUs and taxa (Table S4). Measurable grazing ratios (independent of treatment or sampling date) varied over seven orders of magnitude (Fig. 4), indicating a large range in apparent copepod feeding selection. High ratios were interpreted as suggestive of positive feeding selection, while low ratios were suggestive of negative feeding selection. Highest grazing ratios in this study were observed for large pennate and centric diatoms, fungi, kinetoplastid protozoans, marine invertebrates, heterotrophic nanoflagellates, hypotrich ciliates and unknown eukaryotes, while lowest grazing ratios were observed for oligotrich and choreotrich ciliates, dinoflagellates, autotrophic flagellates, haptophytes including *P. pouchetii* and small diatoms such as *Skeletonema* (Table S4, Part C). Grazing ratios for *P. pouchetii* and *S. marinoi* were almost always low (> 1), potentially indicative of negative selection of these food particles by *Calanus* copepod incubated in mesocosm chambers despite their high relative abundance in the NP and NPSi mesocosms (Fig. 3 and Ray *et al.* 2016). The temporal dynamics of molecular grazing ratios calculated from AmpliconNoise OTU- and CREST taxonomic assignments for each treatment are shown in Fig. S4.

Correlation of intracellular metabolites with metabarcoding results

In total, 274 features were detected in the intracellular metabolite profiles from all seawater samples (Table S5). Canonical correspondence analysis (CCA) with forward selection of peaksum-normalized metabolite profiles from all sampling days and treatments identified

Treatment (Pr(>F)=0.005, F=2.0215, df=3) and Date (Pr(>F)=0.005, F=2.9512, df=1) as significant structuring variables. Visual inspection of hierarchical clustering patterns in metabolite profiles based on chl *a*, phaeophytin and *P. pouchetii* colonial cell abundance dynamics reported previously (Ray *et al*. 2016) for the NP or NPSi mesocosms revealed three distinct stages of the mesocosm succession (Fig. 5). The "early" stage of succession, from 11-17 March, was characterized by the exponential growth of the diatom *S. marinoi*, as indicated by an exponential increase in 0.2 µm-filterable chl *a* (Ray *et al.* 2016) as well as microscopy counts of *S. marinoi* (Fig. 3). The subsequent plateau of chl *a* and concomitant increase in phaeophytin were indicators for the "middle" succession phase, which occurred from 18-23 March. Increasing numbers of *P. pouchetii* colonial cells (an indicator of exponential *P. pouchetii* growth) (FlowCAM measurements, Ray *et al.* 2016) delimited the "late" phase of mesocosms succession, from 24-30 March.

Succession stage was shown to be a highly significant grouping variable for ordination of the NP and NPSi intracellular metabolite profiles by CAPdiscrim analysis ($P = 0.001$ with 1000 permutations), explaining 94-95% of the variation in the principal components of the metabolite data (Fig. 6). Succession phase could also explain 49% of the variation in biomass estimates for microbial eukaryote taxonomic groups based on microscopy counts (available on http://datadryad.org) from NP and NPSi mesocosm seawater (PERMANOVA, df = 2, SS = 4.0224, F = 20.027, Pr(>F) = 0.001). CAPdiscrim analysis furthermore allowed us to identify intracellular metabolites, in particular carbohydrates and carbohydrate derivatives, which could discriminate between the *S. marinoi*-dominated "middle" succession phase and the *P. poucheti*-dominated "late" succession phase during the experiment (Fig. 5). For example, inositol isomers

(e.g Metabolites 159, 173, 174) as well as saccharides (Metabolite 161) correlated significantly with the *S. marinoi* dominated succession phase, while Metabolites 134 and 207 were found to be saccharide-like metabolites whose high concentrations co-occurred significantly with *P. pouchetii* exponential growth. Furthermore, we identified several lipids that were either significantly correlated to the *S. marinoi*-dominated "middle" mesocosm phase, which included fatty acids (e.g. Metabolite 194), or to the *P. pouchetii*-dominated "late" mesocosm phase, which included a terpene (Metabolite 214) and sterols (Metabolites 260, 264, 267). Metabolites associated with the CAPdiscrim-identified mesocosm succession stages are shown in (Fig. 5). Based on microscopy biomass estimates (Table S6), several taxonomic groups were found to be significantly correlated to specific metabolites (Table S7). Weak positive correlation to Metabolites 173 (inositol isomer) and 205 (glucose derivative), for example, were identified for *Skeletonema* biomass estimates in NP mesocosms. Metabolites 203 (1-octadecanol), 207 (unknown) and 209 (unknown) were found to be positively correlated to *Phaeocystis* biomass, while metabolite 262 (cholesterol) was found to be negatively correlated to *Phaeocystis* biomass, in NP mesocosms.

To further test whether *Calanus* grazing ratios on individual prey taxa correlated significantly to metabolites from all mesocosm treatments and Raunefjorden, we performed pairwise correlation analysis of the 274 detected metabolites to *Calanus* grazing ratios calculated from AmpliconNoise OTUs that were collated by fifth-rank taxonomic assignments (Table S8). Due to limited metabarcoding sample numbers and to missing data points for grazing ratios calculated from the metabarcoding data, it was not possible to identify statistically significant correlations between grazing ratio dynamics for individual OTUs/taxa and individual metabolites. Closer inspection of these non-significant results, however, revealed consistent correlative

391 "associations" of grazing ratios with sugar derivatives, saccharides, amino acids and their
392 derivatives, and fatty acids, thus implicating these metabolite classes in *Calanus* grazing
393 selection (Table S8). Metabolites 159, 173 and 194, for example, which were significantly
394 correlated with the *S. marinoi*-dominated "middle" mesocosm metabolic phase, were found to be
395 negatively associated with the grazing ratios for Bacillariophyta in the diatom-dominated NPSi
396 mesocosm treatment (Kendall's *tau* = -0.733, -0.87 and -0.733, respectively) (Table S8). For the
397 NP mesocosms, which experienced strongest dominance by *P. pouchetii* during the "late"
398 mesocosm metabolic phase, metabolites 134, 214, 260, 264 and 267 were found to be positively
399 associated with *Calanus* grazing ratios on Phaeocystales (Kendall's *tau* = 0.966, 0.690, 0.828,
400 0.552 and 0.552 respectively). Comprehensive results from correlation analysis of grazing ratios
401 with metabolites can be found in Table S8.
402
403 **Discussion**
404 **Selective feeding by *Calanus* copepods**
405 Our molecular characterization of *Calanus* sp. grazing during a seawater mesocosm experiment
406 provides new evidence for dynamic and discriminate feeding selection by this copepod in mixed
407 natural microbial plankton assemblages. Although *Calanus* sp. copepods may generate feeding
408 currents for passive grazing, our findings support numerous empirical studies demonstrating
409 clear selectivity in copepod feeding behavior (Huntley 1988; Meyer-Harms *et al.* 1999;
410 Nejstgaard *et al.* 1997; 2008; Barofsky *et al.* 2010, and references therein) in mixed natural
411 assemblages of microbial plankton. Nutrient manipulation of mesocosms to promote a *P.*
412 *pouchetii* bloom further indicated that the increase in prey particle size range caused by the
413 formation and growth of *P. pouchetii* colonies during bloom development does not increase the
414 ingestion of *P. pouchetii* by *Calanus*, negating size selection as a significant determinant of

copepod feeding (Nejstgaard *et al*. 2007). On the contrary, our molecular diet analysis provides direct evidence that *P. pouchetii* does not contribute significantly to *Calanus* copepod diet in a mesocosm setting (Nejstgaard *et al.* 2006; Ray *et al*. 2016). The timing of the bloom-like growth of *P. pouchetii* in this experiment was such that peak abundance occurred around the time when the experiment was stopped. The complete sample set does therefore not include any samples from a *P. pouchetii* "bloom decline" stage, precluding our ability to test the hypothesis of Estep *et al.* (1990) that ageing or senescent *P. pouchetii* colonies, either by merit of physical degradation or changes in chemical properties, are more readily consumed by *Calanus* copepods than younger colonies.

Our molecular grazing ratio results are in accord with idealized food web models and experimental observations in which mesozooplankton (copepods) have been shown to prefer ciliates and large diatoms over small autotrophs as a food source (e.g. Kleppel *et al.* 1991; Stoecker & Capuzzo 1990; Ohman & Runge 1994; Nejstgaard *et al.* 1997, 2001b; Calbet & Saiz 2005; Yang *et al.* 2009; Fileman *et al.* 2010). The low grazing selection ratios that we observed for dinoflagellate taxa including *Gyrodinium*, the most abundant genus observed in our experiment (Stoecker *et al*. 2015, Ray *et al*. 2016), were more surprising as previous studies have shown high ingestion rates of dinoflagellates by calanoid copepods (Levinsen *et al.* 2000; Batten *et al.* 2001; Olson *et al.* 2006). For example, Fileman *et al.* (2010) observed high clearance rates by *C. helgolandicus* on *Gyrodinium fusiforme*/*spirale* in June in the English Channel; at this time large heterotrophic dinoflagellates dominated the microzooplankton biomass and large choreotrich ciliates were scarce. However, in most grazing experiments, choreotrich ciliates accounted for a larger proportion of the microzooplankton carbon ingested than did heterotrophic

dinoflagellates (Fileman *et al.* 2010). Based on our grazing proxy, our results suggest that *Gyrodinium* spp. dinoflagellates were not prefered prey during the mesocosm experiment. One explanation for this finding is active avoidance of *Calanus* predation by dinoflagellates (Granéli *et al.* 1993; Nejstgaard et al. 1997; 2001a; Verity 2010). Alternatively, ingested dinoflagellates might be rapidly digested in the *Calanus* digestive tract (e.g. Sullivan 2011), however we did not observe consistently low grazing ratios for all naked autotroph and protist taxa present in the grazing ratio data (Table S4), nor are we aware of studies demonstrating taxon-dependent differential prey DNA digestion rates. An additional surprising finding of our grazing ratio proxy is the apparent positive selection by *Calanus* copepods on fungi. Dikarya were highly grazed in all treatments and Raunefjorden throughout the experiment (Fig. S4). Inspection of taxonomic classification of these taxa revealed primarily pezizomycotina (Ascomycota) and agaricomycetes (Basidiomycota). The high relative presence of these fungal taxa in *Calanus* gut content may suggest grazing on fungal hyphae or spores, but may also suggest the consumption of detrital material by copepods (although see Paffenhöffer & Strickland 1970), as these fungal groups are known saprophytes of marine macroalgae (Richards *et al.* 2012). It should be further noted that the apparent gut content of harvested copepods likely reflects prey particles consumed only in the time period (< 1 hour) immediately prior to sampling, as prey DNA digestion (Troedsson *et al.* 2009) and rapid copepod gut passage (Nejstgaard *et al*. 2003) would certainly hinder detection of prey particles cumulatively consumed during the entire mesocosm incubation period. The metabarcoding data from copepod gut content samples thus represents a snapshot of grazing activity by *Calanus* copepods incubated in mesocosm grazing chambers. In order to ensure that the seawater microbial plankton diversity would provide a concurrent diversity "scaffold" upon which to assess prey selection by feeding copepods, seawater samples for

metabarcoding analysis were collected from mesocosms and Raunefjorden immediately prior to copepod harvesting.

The interpretation of metabarcoding results as an indication of positive or negative grazing selection by *Calanus* on specific OTUs or taxa requires the implicit assumption that metabarcoding abundances correspond to biologically meaningful quantities, i.e. prey abundances. The potential sources of PCR and target gene copy number bias that can limit the quantitative power of metabarcoding approaches in microbial ecology have been described in other publications (Richards & Bass 2005; Potvin & Lovejoy 2009; Amend *et al*. 2010; Stoeck *et al.* 2010; Deagle *et al.* 2010; 2013; Gong *et al*. 2013). Cognizant of these cautions, we have utilized grazing ratios rather than absolute abundances, thus normalizing read abundance data for some of the bias inherent in PCR-based sequencing library preparation. We demonstrate correspondence between metabarcoding analysis, microscopy counts and qPCR estimations for key mesocosm taxa, namely *P. pouchetii* and *S. marinoi*. The semi-quantitative interpretation of metabarcoding results for individual OTUs or taxa (D'Amore *et al.* 2016; Lanzén *et al*. 2016) has thus generated a numerical proxy by which grazing selection may be evaluated, and which has provided a useful tool for correlative analysis of copepod feeding selection with plankton metabolic succession. Contemporary calls for an integrated approach to marine biodiversity assessment and baseline establishment include the contribution of molecular analysis (Duffy *et al.* 2013), thus critical evaluations of the quantitative power of, for example, metabarcoding data to investigate trophic interactions (King *et al.* 2008), are both timely and relevant.

Metabarcoding analysis marker selection

Sequence read length limitations of earlier generation sequencing technologies restricted the choice of SSU rRNA hypervariable target region to those regions (typically V9 for eukaryotes, ~200 bp) representing the best compromise between phylogenetic information density and shortest possible target fragment length (Amaral-Zettler *et al.* 2009; Brown *et al.* 2009; Stoeck *et al.* 2009; Edgcomb *et al.* 2011). Later improvements in sequencing chemistry, however, have facilitated sequencing of longer target regions, thus increasing the possibility for exploration of microbial diversity using other hypervariable regions within the SSU rRNA gene (Chariton *et al.* 2010; Stoeck *et al.* 2010; Behnke *et al.* 2011; Monchy *et al.* 2011; Lanzén *et al.* 2016). Molecular analysis of trophic interactions, however, must also assume considerable prey DNA degradation inside the host gut, which reduces the informational advantage obtained with longer sequence reads (Troedsson *et al.* 2009). The diversity of prey organisms observed in copepod gut content samples in this study (Fig. 1) confirms that the V7 region of the SSU rRNA gene targeted in this study (Hadziavdic *et al*. 2014) provides a satisfactory compromise between fragment length (260-360 bp) and phylogenetic/taxonomic resolution (Table S1) for broad exploration of microbial eukaryote plankton diversity in the context of copepod grazing selection.

Putative protist symbionts of *Calanus*

The most abundant taxa identified among classified reads and OTUs in the copepod gut content sequence data had highest similarity to the taxonomically ambiguous *Paradinium* (Shields 1994; Carman & Dobbs 1997; Skovgaard & Daugbjerg 2008), the dinoflagellate groups Blastodinia and Syndinia (Shields 1994), Apicomplexa (Rueckert *et al.* 2011), and the apostome ciliate clade Oligohymenophorea (Carman and Dobbs 1997; Prokopowicz *et al.* 2010; Guo *et al.* 2012;

Chantangsi *et al.* 2013). Symbioses between *Calanus* and commensals and/or parasites would be characterized by our proxy as having high grazing ratios despite the fact that *Calanus* had not consumed these organisms *per se*. We have therefore not drawn conclusions about *Calanus* grazing activity on these taxa. Furthermore, the nature of the biological relationship of these taxa with the *Calanus* copepods in which they were detected falls outside the scope of this study, yet raises questions about the incidence of *Calanus* parasitism and the effect of parasitism on copepod health, reproduction and feeding behavior (Skovgaard & Saiz 2006; Cirtwill *et al*. 2015; Worden *et al*. 2015). Future studies of *Calanus* feeding using state-of-the-art high-throughput sequencing technology might increase the sequencing depth of *Calanus* gut content sample coverage to improve the sensitivity of detection of prey organisms even in the presence of symbioses. Alternatively, prey enrichment strategies such as blocking PCR oligos (Troedsson *et al.* 2008b; Hu *et al*. 2014) against dominant suspected symbiont sequences, DHPLC-PCR (Troedsson *et al.* 2008a,b; Olsen *et al.* 2012; 2014) and/or restriction enzyme treatment (Maloy *et al.* 2013) might also be employed to overcome symbiont phylogenetic signal. Amplicon libraries were not generated from starved copepods in this study, which would have enabled a more conclusive assessment of prey versus symbiont. The sequencing strategy employed in this study to investigate *Calanus* grazing choice has nonetheless uncovered a high diversity of *Calanus*-associated microbial eukaryotes that suggests a considerable degree of endemic symbioses in copepods in accord with previous observations (Skovgaard & Saiz 2006; Rueckert *et al.* 2011; Bickel *et al.* 2012*;* Lima-Mendez *et al.* 2015).

Role of Metabolites in Grazing Selection by *Calanus*

The importance of carbohydrates and lipids was a feature common to all correlation analyses

(metabolite succession phase, biomass of taxonomic groups, molecular grazing ratios) performed in this study, implicating the importance of these metabolite classes in the trophodynamics of *Calanus* feeding. For example, inositol-related peaks (Metabolites 159, 173 and 174) were positively correlated with the *S. marinoi*-dominated "middle" metabolite succession phase as well as *Skeletonema* biomass estimates, but negatively associated to *Calanus* grazing ratios on diatoms, in the NPSi treatment. Using a qPCR-based molecular grazing proxy, we have previously shown that *Calanus* grazing selection on *S. marinoi* was low during exponential growth of *S. marinoi*, but increased with diatom bloom senescence (Ray *et al.* 2016). Indeed, grazing ratios for diatoms, albeit generally low, increased toward the end of the mesocosm experiment in the NP and NPSi treatments in this study (Fig. S5, "Bacillariophyta"), suggesting that copepod grazing on diatoms only increased as the diatom bloom declined during the "late" metabolite succession phase. These results further suggest that exponential-phase *S. marinoi* is not preferred prey for *Calanus*, and that *Calanus* grazing selection may rather be linked to the developmental status of the greater prey community that is reflected in its chemical composition, as previously demonstrated (Barofsky *et al*. 2010; Ray *et al.* 2016).

The complexity of plankton succession in natural assemblages presents an ongoing challenge of our ability to identify taxon-associated metabolites and their role(s) in grazing selection by copepods. For example, Metabolite 260, a steroid, was significantly positively correlated to the "late" metabolome succession phase of the NP and NPSi treatments, however it was not significantly correlated to microscopy-based biomass estimates for *P. pouchetii* (Table S7), which was a dominant phytoplankton taxon during this phase. This suggests correspondence of this metabolite to the "late" succession phase itself, rather than to one specific taxon. Our

additional observation that this metabolite was positively associated with *Calanus* grazing ratios on *P. pouchetii* in the NP treatment (Table S8) is therefore likely anecdotal, i.e. that *Calanus* simultaneously consumed *P. pouchetii* colonies or colony fragments that were physically associated with other prey particles such as diatoms (Smetacek *et al*. 2002), whose lipids and carbohydrates may have become more bioavailable to *Calanus* due to diatom bloom senescence during the "late" succession phase. Indeed, Metabolite 207 (galactosylglycerol) was weakly positively correlated with *P. pouchetii* biomass estimates (Table S7), but also positively associated with molecular grazing ratios for diatoms (Table S8). The positive association between Metabolite 207 and *P. pouchetii* biomass may therefore indicate that the underlying biological interaction was in fact due to the co-occurring *S. marinoi* bloom decline rather than to the increase in *P. pouchetii* biomass *per se*. In concert, these results highlight the biological and chemical complexity of the seawater mesocosm plankton communities.

Conclusions

In order to generate new knowledge about copepod grazing behavior, we have implemented a proxy for *Calanus* feeding selection based on metabarcoding analysis of eukaryotic microbial plankton communities. This proxy has revealed a diverse prey landscape, and allowed us to identify associations between copepod feeding selection dynamics and metabolic signatures of co-occurring microbial plankton communities. Our correlative approach has provided consistent indication of the importance of carbohydrates and lipids during *Calanus* feeding selection. In line with the findings of previous studies, the combined metabarcoding and metabolomic analyses therefore suggest that prey selection by copepods is determined by the developmental status of the diverse prey community, rather than by relative abundances of individual prey taxa. We have

demonstrated that the combined analysis of high-resolution biological and chemical factors can provide new information about copepod feeding selecting in dynamic and complex prey assemblages. This knowledge improves our understanding of the efficiency of the marine food web and increases our ability to predict responses to perturbation and changing climate on a temporal and spatial scale relevant for evolutionary forces on plankton.

Acknowledgments

This study was funded by the Research Council of Norway (RCN) research project "A novel cross-disciplinary approach to solve an old enigma: the food-web transfer of the mass-blooming phytoplankter *Phaeocystis pouchetii*" (Phaeonigma, project number 204479/F20). Additional support was received from the European Research Council Advanced Grant ERC-AG-LS8 "Microbial Network Organisation" (MINOS, project number 250254), the German Research Foundation within the framework of the CRC1127 ChemBioSys and the RCN project "Processes and players in Arctic marine pelagic food webs - biogeochemistry, environment and climate change" (MicroPolar, project number 225956/E10). JLR received Short Term Scientific Mission travel support through the COST-EU funding program ES1103-Microbial ecology & the earth system: collaborating for insight and success with the new generation sequencing tools (COST-STSM-ES1103-16732 and COST□STSM□ES1103□21370).

References

Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS One 4: e6372

Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: does read abundance count? Mol Ecol 19: 5555-5565

601 Anderson MJ, Willis TJ (2003) Canonical analysis of principal coordinates: a useful method of
602 constrained ordination for ecology. Ecol 84: 511-525

603 Anderson MJ (2004) CAP: a FORTRAN computer program for canonical analysis of principal
604 coordinates. Department of Statistics, University of Auckland, New Zealand

605 Ausloos P, Clifton CL, Lias SG, Mikaya AI, Stein SE, Tchekhovskoi DV, Sparkman OD, Zaikin
606 V, Zhu D (1999) The critical evaluation of a comprehensive mass spectral library. J Am Soc
607 Mass Spec 10: 287-299

608 Barofsky A, Simonelli P, Vidoudez C, Troedsson C, Nejstgaard JC, Jakobsen HH, Pohnert G
609 (2010) Growth phase of the diatom *Skeletonema marinoi* influences the metabolic profile of the
610 cells and the selective feeding of the copepod *Calanus* spp. J Plankt Res 32: 263-272

611 Batten SD, Fileman ES, Halvorsen E (2001) The contribution of microzooplankton to the diet of
612 mesozooplankton in an upwelling filament off the north west coast of Spain. Prog Oceanogr 51:
613 385-398

614 Behnke A, Engel M, Christen R, Nebel M, Klein RR, Stoeck T (2011) Depicting more accurate
615 pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA
616 gene regions. Environ Microbiol 13: 340-349

617 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful
618 approach to multiple testing. J Roy Stat Soc B 57: 289-300

619 Bickel SL, Tang KW, Grossart H-P (2012) Ciliate epibionts associated with crustacean
620 zooplankton in German lakes: Distribution, motility and bacterivory. Front Microbiol 3: 27-54

621 Brown MV, Philip GK, Bunge JA, Smith MC, Bissett A, Lauro FM, Fuhrman JA, Donachie SP
622 (2009) Microbial community structure in the North Pacific ocean. ISME J 3: 1374-1386

623 Calbet A, Saiz E (2005) The ciliate-copepod link in marine ecosystems. Aquat Microb Ecol 38:
624 157-167

625 Carman KR, Dobbs FC (1997) Epibiotic microorganisms on copepods and other marine
626 crustaceans. Microscop Res Tech 37: 116-135

627 Chantangsi C, Lynn DH, Rueckert S, Prokopowicz AJ, Panha S, Leander BS (2013) *Fusiforma*
628 *themisticola* n. gen., n. sp., a new genus and species of apostome ciliate infecting the hyperiid
629 amphipod *Themisto libellula* in the Canadian Beaufort Sea (Arctic Ocean), and establishment of
630 the Pseudocolliniidae (Ciliophora, Apostomatia). Protist 164: 793-810

631 Chariton AA, Court LN, Hartley DM, Colloff MJ, Hardy CM (2010) Ecological assessment of
632 estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. Front Ecol Environ 8: 233-

238

Cirtwill AR, Stouffer DB (2015) Concomitant predation on parasites is highly variable but constrains the ways in which parasites contribute to food web structure. J Anim Ecol 84: 734-744

Cowles TJ, Olson RJ, Chisholm SW (1988) Food selection by copepods: discrimination on the basis of food quality. Mar Biol 100: 41-49

D'Amore R, Ijaz UZ, Schirmer M, Kenny JG, Gregory R, Darby AC, Shakya M, Podar M, Quince C, Hall N (2016) A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. BMC Genomics 17: 55

Deagle BE, Chiaradia A, McInnes J, Jarman SN (2010) Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? Conserv Genet 11: 2039-2048

Deagle BE, Thomas AC, Shaffer AK, Trites AW, Jarman SN (2013) Quantifying sequence proportions in a DNA□based diet study using Ion Torrent amplicon sequencing: which counts count? Mol Ecol Res 13: 620-633

Duffy JE, Amaral-Zettler LA, Fautin DG, Paulay G, Rynearson TA, Sosik HM, Stachowicz JJ (2013) Envisioning a Marine Biodiversity Observation Network. 63: 350-361.

Edgar RC (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Meth 10: 996-998

Edgcomb V, Orsi W, Bunge J, Jeon S, Christen R, Leslin C, Holder M, Taylor GT, Suarez P, Varela R, Epstein S (2011) Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. ISME J 5: 1344-1356

Estep KW, Nejstgaard JC, Skjoldal HR, Rey F (1990) Predation by copepods upon natural populations of *Phaeocystis pouchetii* as a function of the physiological state of the prey. Mar Ecol Prog Ser 67: 235-249

Fileman E, Petropavlovsky A, Harris R (2010) Grazing by the copepods *Calanus helgolandicus* and *Acartia clausi* on the protozooplankton community at station L4 in the Western English Channel. J Plankt Res 32: 709-724

Gong J, Dong J, Liu X, Massana R (2013) Extremely high copy numbers and polymorphisms of the rDNA operon estimated from single cell analysis of oligotrich and peritrich ciliates. Protist 164: 369-379

Granéli E, Olsson P, Carlsson P, Granéli W, Christer N (1993) Weak 'top-down'control of dinoflagellate growth in the coastal Skagerrak. J Plankt Res 15: 213-237

664 Guo Z, Liu S, Hu S, Li T, Huang Y, Liu G, Zhang H, Lin S (2012) Prevalent ciliate symbiosis on
665 copepods: high genetic diversity and wide distribution detected using small subunit ribosomal
666 RNA gene. PLoS One 7: e44847

667 Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, Troedsson C (2014)
668 Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. PLoS
669 One 9: e87624

670 Hu S, Guo Z, Li T, Carpenter EJ, Liu S, Lin S (2014) Detecting *in situ* copepod diet diversity
671 using molecular technique: development of a copepod/symbiotic ciliate-excluding eukaryote-
672 inclusive PCR protocol. PLoS One 9: e103528.

673 Huntley M (1988) Feeding biology of *Calanus*: a new perspective. Hydrobiol 167: 83-99

674 Irigoien X (1998) Gut clearance rate constant, temperature and initial gut contents: a review. J
675 Plankt Res 20: 997-1003

676 King RA, Read DS, Traugott M, Symondson WO (2008) Molecular analysis of predation: a
677 review of best practice for DNA□based approaches. Mol Ecol 17: 947-963

678 Kiørboe T, Saiz E, Viitasalo M (1996) Prey switching behaviour in the planktonic copepod
679 *Acartia tonsa*. Mar Ecol Prog Ser 143: 65-75

680 Kiørboe T (2011) How zooplankton feed: mechanisms, traits and trade-offs. Biol Rev 86: 311-
681 340

682 Kleppel GS, Holliday DV, Pieper RE (1991) Trophic interactions between copepods and
683 microplankton: a question about the role of diatoms. Limnol Oceanogr 36: 172-178

684 Kleppel GS (1993) On the diets of calanoid copepods. Mar Ecol Prog Ser 99: 183-195

685 Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2012)
686 Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation
687 sequencing-based diversity studies. Nucl Acids Res gks808

688 Kuhlisch C, Pohnert G (2015) Metabolomics in chemical ecology. Nat Prod Rep 32: 937-955

689 Landry MR (2002) Integrating classical and microbial food web concepts: evolving views from
690 the open-ocean tropical Pacific. Hydrobiologia 480: 29-39

691 Lanzén A, Jørgensen SL, Huson DH, Gorfer M, Grindhaug SH, Jonassen I, Øvreås L, Urich T
692 (2012) CREST–Classification resources for environmental sequence tags. PLoS One 7: e49334

693 Lanzén A, Lekang K, Jonassen I, Thompson EM, Troedsson C (2016) High□throughput

metabarcoding of eukaryotic diversity for environmental monitoring of offshore oil drilling activities. Mol Ecol (*in press*): 10.1111/mec.13761

Leung TLF, Polin R (2008) Parasitism, commensalism, and mutualism: Exploring the many shades of symbioses. Life & Environ 58: 107-115.

Levinsen H, Turner JT, Nielsen TG, Hansen BW (2000) On the trophic coupling between protists and copepods in arctic marine ecosystems. Mar Ecol Prog Ser 204: 65-77

Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, Chaffron S, Ignacio-Espinosa JC, Roux S, Vincent F, Bittner L, Darzi Y, Wang J, Audic S, Berline L, Bontempi G, Cabello AM, Coppola L, Cornejo-Castillo FM, d'Ovidio F, De Meester L, Ferrera I, Garet-Delmas M-J, Guidi L, Lara E, Pesant S, Royo-Llonch M, Salazar G, Sánchez P, Sebastian M, Souffreau C, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Tara Oceans coordinators, Gorsky G, Not F, Ogata H, Speich S, Stemmann L, Weissenbach J, Wincker P, Acinas SG, Sunagawa S, Bork P, Sullivan MB, Karsenti E, Bowler C, de Vargas C, Raes J (2015) Determinants of community structure in the global plankton interactome. *Science* 348: 1262073

Maloy AP, Culloty SC, Slater JW (2013) Dietary analysis of small planktonic consumers: a case study with marine bivalve larvae. J Plankt Res: fbt027

Meyer-Harms B, Irigoien X, Head R, Harris R (1999) Selective feeding on natural phytoplankton by *Calanus finmarchicus* before, during, and after the 1997 spring bloom in the Norwegian Sea. Limnol Oceanogr 44: 154-165

Monchy S, Sanciu G, Jobard M, Rasconi S, Gerphagnon M, Chabé M, Cian A, Meloni D, Niquil N, Christaki U, Viscogliosi E (2011) Exploring and quantifying fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing. Environ Microbiol 13: 1433-1453

Nejstgaard JC, Gismervik I, Solberg PT (1997) Feeding and reproduction by *Calanus finmarchicus*, and microzooplankton grazing during mesocosm blooms of diatoms and the coccolithophore *Emiliania huxleyi*. Mar Ecol Prog Ser 147: 197-217

Nejstgaard JC, Naustvoll L-J, Sazhin A (2001a) Correcting for underestimation of microzooplankton grazing in bottle incubation experiments with mesozooplankton. Mar Ecol Prog Ser 221: 59-75

Nejstgaard JC, Hygym B, Naustoll L-J, Båmstedt U (2001b) Zooplankton growth, diet and reproductive success compared in simultaneous diatom- and flagellate-microzooplankton dominated plankton blooms. Mar Ecol Prog Ser 221: 77-91

Nejstgaard JC, Frischer ME, Raule CL, Gruebel R, Kohlberg KE, Verity PG (2003) Molecular

detection of algal prey in copepod guts and faecal pellets. Limnol Oceanogr Meth 1: 29-38

Nejstgaard JC, Frischer ME, Verity PG, Anderson JT, Jacobsen A, Zirbel MJ, Larsen A, Martínez-Martínez J, Sazhin AF, Walters T, Bronk DA (2006) Plankton development and trophic transfer in seawater enclosures with nutrients and *Phaeocystis pouchetii* added. Mar Ecol Prog Ser 321: 99-121

Nejstgaard JC, Tang KW, Steinke M, Dutz J, Koski M, Antajan E, Long JD (2007) Zooplankton grazing on *Phaeocystis*: a quantitative review and future challenges. Biogeochem 83: 147-172

Nejstgaard JC, Frischer ME, Simonelli P, Troedsson C, Brakel M, Adiyaman F, Sazhin AF, Artigas LF (2008) Quantitative PCR to estimate copepod feeding. Mar Biol 153: 565-577

Ohman MD, Runge JA (1994) Sustained fecundity when phytoplankton resources are in short supply: Omnivory by *Calanus finmarchicus* in the Gulf of St. Lawrence. Limnol Oceanogr 39: 21-36

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2015) vegan: Community Ecology Package. R package version 2.3-0. http://CRAN.R-project.org/package=vegan

Olsen BR, Dahlgren K, Schander C, Båmstedt U, Rapp HT, Troedsson C (2012) PCR–DHPLC assay for the identification of predator–prey interactions. J Plankt Res 27: fbr110

Olsen BR, Troedsson C, Hadziavdic K, Pedersen RB, Rapp HT (2014) A molecular gut content study of *Themisto abyssorum* (Amphipoda) from Arctic hydrothermal vent and cold seep systems. Mol Ecol 23: 3877-3889

Olson MB, Lessard EJ, Wong CHJ, Bernhardt MJ (2006) Copepod feeding selectivity on microplankton, including the toxigenic diatoms Pseudo-nitzschia spp. in the coastal northwest Pacific. Mar Ecol Prog Ser 326: 207-220

Pepin P, Colbourne E, Maillet G. (2011) Seasonal patterns in zooplankton community structure on the Newfoundland and Labrador Shelf. Prog Oceanogr 91: 274-285

Pitta P, Nejstgaard JC, Tsagaraki TM, Zervoudaki S, Egge JK, Frangoulis C, Lagaria A, Magiopoulos I, Psarra S, Sandaa R-A, Skjoldal EF, Tanaka T, Thyrhaug R, Thingstad TF (2016) Confirming the "Rapid phosphorous transfer from microorganisms to mesozooplankton in the Eastern Mediterranean Sea" scenario through a mesocosm experiment. J Plankt Res, *in press*

Pompanon F, Deagle BE, Symondson WO, Brown DS, Jarman SN, Taberlet P (2012) Who is eating what: diet assessment using next generation sequencing. Mol Ecol 21: 1931-1950

Potvin M, Lovejoy C (2009) PCR□based diversity estimates of artificial and environmental 18S

rRNA gene libraries. J Euk Microbiol 56: 174-181

Poulet SA, Marsot P (1978) Chemosensory grazing by marine calanoid copepods (Arthropoda: Crustacea). Science 200: 1403-1405.

Prokopowicz AJ, Rueckert S, Leander BS, Michaud J, Fortier L (2010) Parasitic infection of the hyperiid amphipod *Themisto libellula* in the Canadian Beaufort Sea (Arctic Ocean), with a description of *Ganymedes themistos* sp. n. (Apicomplexa, Eugregarinorida). Polar Biol 33: 1339-1350

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucl Acids Res 41: D590-D596

Quince C, Lanzén A, Curtis TP, Davenport RJ, Hall N, Head IM, Read LF, Sloan WT (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. Nature Meth 6: 639-641

Quince C, Lanzén A, Davenport RJ, Turnbaugh PJ (2011). Removing noise from pyrosequenced amplicons. BMC Bioinform 12: 38

Paffenhöfer GA, Strickland JDH (1970) A note on the feeding of *Calanus helgolandicus* on detritus. *Mar Biol* 5: 97-99

R Development Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria

Ray JL, Skaar KS, Simonelli PS, Larsen A, Sazhin A, Jakobsen HH, Nejstgaard JC, Troedsson C (2016) Molecular gut content analysis demonstrates that *Calanus* grazing on *Phaeocystis pouchetii* and *Skeletonema marinoi* is sensitive to bloom phase but not prey density. Mar Ecol Prog Ser 542: 63-77

Richards TA, Bass D (2005) Molecular screening of free-living microbial eukaryotes: diversity and distribution using a meta-analysis. Curr Op Microbiol 8: 240-252

Richards TA, Jones MDM, Leonard G, Bass D (2012) Marine fungi: their ecology and molecular diversity. Annu Rev Mar Sci 4: 495-522

Rueckert S, Simdyanov TG, Aleoshin VV, Leander BS (2011) Identification of a divergent environmental DNA sequence clade using the phylogeny of gregarine parasites (Apicomplexa) from crustacean hosts. PLoS One 6: e18163d quality and population control. Aquat Microb Ecol 7: 197-223

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,

791 Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF
792 (2009) Introducing mothur: Open-source, platform-independent, community-supported software
793 for describing and comparing microbial communities. Appl Environ Microbiol 75: 7537-7541

794 Sherr E, Sherr B (1988) Role of microbes in pelagic food webs: A revised concept. Limnol
795 Oceanogr 33: 1225-1227

796 Shields JD (1994). The parasitic dinoflagellates of marine crustaceans. Annu Rev Fish Diseases
797 4: 241-271

798 Simonelli P, Troedsson C, Nejstgaard JC, Zech K, Larsen JB, Frischer ME (2009) Evaluation of
799 DNA extraction and handling procedures for PCR-based copepod feeding studies. J Plankt Res
800 31: 1465-1474

801 Skovgaard A, Saiz E (2006) Seasonal occurrence and role of protistan parasites in coastal marine
802 zooplankton. Mar Ecol Prog Ser 327: 37-49

803 Skovgaard A, Daugbjerg N (2008) Identity and systematic position of *Paradinium poucheti* and
804 other *Paradinium*-like parasites of marine copepods based on morphology and nuclear-encoded
805 SSU rDNA. Protist 159: 401-413

806 Smetacek V, Klaas C, Menden-Deuer S, Rynearson TA (2002) Mesoscale distribution of
807 dominant diatom species relative to the hydrographical field along the Antarctic Polar Front.
808 Deep Sea Res II 49: 3835-3848

809 Stoeck T, Behnke A, Christen R, Amaral-Zettler L, Rodriguez-Mora MJ, Chistoserdov A, Orsi
810 W, Edgcomb VP (2009) Massively parallel tag sequencing reveals the complexity of anaerobic
811 marine protistan communities. BMC Biol 7: 72

812 Stoeck T, Bass D, Nebel M, Christen R, Jones MD, Breiner HW, Richards TA (2010) Multiple
813 marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic
814 community in marine anoxic water. Mol Ecol 19: 21-31

815 Stoecker DK, Capuzzo JM (1990) Predation on protozoa: its importance to zooplankton. J Plankt
816 Res 12: 891-908

817 Stoecker DK, Nejstgaard JC, Madhusoodhanan R, Pohnert G, Wolfram S, Jakobsen HH, Šulčius
818 S, Larsen A (2015) Underestimation of microzooplankton grazing in dilution experiments due to
819 inhibition of phytoplankton growth. Limnol Oceanogr 60: 1426-1438

820 Sullivan LJ (2010) Gut evacuation of larval *Mnemiopsis leidyi* A. Agassiz (Ctenophora, Lobata).
821 J Plankt Res 32: 69-74

822 Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation

biodiversity assessment using DNA metabarcoding. Mol Ecol 21: 2045-2050

Troedsson C, Lee RF, Stokes V, Walters TL, Simonelli P, Frischer ME (2008a) Development of a denaturing high-performance liquid chromatography method for detection of protist parasites of metazoans. Appl Environ Microbiol 74: 4336-4345

Troedsson C, Lee RF, Walters T, Stokes V, Brinkley K, Naegele V, Frischer ME (2008b) Detection and discovery of crustacean parasites in blue crabs (*Callinectes sapidus*) by using 18S rRNA gene-targeted denaturing high-performance liquid chromatography. Appl Environ Microbiol 74: 4346-4353

Troedsson C, Simonelli P, Nägele V, Nejstgaard JC, Frischer ME (2009) Quantification of copepod gut content by differential length amplification quantitative PCR (dla-qPCR). Mar Biol 156: 253-259

Verity PG (2010) Expansion of potentially harmful algal taxa in a Georgia Estuary (USA). Harmful Algae 9: 144-152

Vestheim H, Jarman SN (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. Front Zool 5: 12

Vidoudez C, Pohnert G (2012) Comparative metabolomics of the diatom *Skeletonema marinoi* in different growth phases. Metabolomics 8: 654-669

Wagner C, Sefkow M, Kopka J (2003) Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. Phytochem 62: 887-900

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer, New York

Woodson CB, Webster DR, Weissburg MJ, Yen J (2007) Cue hierarchy and foraging in calanoid copepods: ecological implications of oceanographic structure. Mar Ecol Prog Ser 330: 163-177

Worden AZ, Follows MJ, Giovannoni SJ, Wilken S, Zimmerman AE, Keeling PJ (2015) Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. Science 347: 736-745

Yang EU, Kang H-K, Yoo S, Hyun J-H (2009) Contribution of auto- and heterotrophic protozoa to the diet of copepods in the Ulleung Basin, East Sea/Japan Sea. J Plankt Res 31: 647-659

Data Accessibility

Pyrosequencing flowgram files (*.sff) have been submitted to the NCBI Sequence Read Project with accession number SRP076974. The AmpliconNoise OTU table, the CREST taxonomic assignment table are available from the Dryad Digital Repository http://dx.doi.org/10.5061/dryad.6tn7c.

Author Contributions

JLR, JA, PS, JCN, AL, DS, MF, GP and CT designed the research. AL and JCN coordinated the mesocosm experiment. JLR, JA, PS, DS, AS, GP, MF and CT collected samples. JLR, JA, KSS, DS and AS performed assay optimization and laboratory processing of samples. JLR, JA, UI and CQ analyzed the data. JLR, JA, PS, JCN, AL, DS, MF, GP and CT wrote the manuscript.

Table and Figure Captions

Table 1. Primers and blocking oligonucletide probes used in this study.

Table 2. PERMANOVA analysis of metabarcoding diversity from OTU clustering (AmpliconNoise) or taxonomic classification (CREST). Significant *P*-values are indicated in bold. Explanatory variable levels tested were: Sample Type - seawater vs. copepod gut content; Experiment - mesocosm vs. fjord; Treatment - Control, NP, NPSi or Raunefjorden; Date - 11, 17, 21, 24, 28 or 30 March 2012.

Table S1. *In silico* coverage of Silva SSU r121 RefNR sequence collection using TestPrime v2.0 (Klindworth *et al.* 2012) with primers F-1183mod and R-1443mod and allowing 0, 1, or 2 mismatches between primers and database sequence. Numbers in parenthesis indicate percentage

of the r121 RefNR sequence collection detected *in silico* for each number of allowed mismatches.

Table S2. Summary metrics from analysis of metabarcoding results from mesocosm seawater (*n* = 24) and copepod samples (*n* = 24) analyzed using AmpliconNoise for OTU clustering or CREST for taxonomic assignment.

Table S3. Read distribution per sample and beta diversity estimates from metabarcoding analysis.

Table S4. Grazing selection by *Calanus* as identified through metabarcoding analysis of copepod gut content and co-occuring seawater plankton assemblages. Molecular grazing ratios (ratio of each OTU/taxon in copepod gut content to their abundance in seawater) were calculated for (A) individual AmpliconNoise OTUs, (B) AmpliconNoise OTUs grouped by fifth-level (or highest available) CREST taxonomic assignments, (C) Highest-level CREST taxonomic assignment ranks and (D) CREST taxonomic assignments grouped by fifth-level (or highest available) taxonomic ranks. Only finite, non-zero values are shown.

Table S5. Metabolome profiles. (A) GC-MS peak identifications. Metabolite, peak number; Ion, ionization energy in *e*V; RT, retention time; Identification, certainty is expressed as the reverse match factor of each metabolite when searched against the National Institute of Standards and Technology chemistry database: from 750 to 850 (?), 650 to 750 (??), or 550 to 650 (???). (B) Peaksum-normalized metabolite concentrations in each mesocosm on all sampling days. Top column names indicate <*treatment_sampling date*>. Bottom column names indicate mesocosm number. (C) Mean metabolite concentration per treatment on each sampling date. Samples (rownames) are indicated as <*treatment_sampling date*>.

899 **Table S6**. Biomass estimates for major eukaryotic plankton taxonomic groups in Control, NP
900 and NPSi mesocosms. Biomass estimates were calculated from microscopy counts according to
901 previously described methods (Ray *et al.* 2016).

902 **Table S7**. Metabolites with significant correlation to biomass estimates for major microbial
903 eukaryote plankton taxonomic groups present in the NP and NPSi seawater mesocosms. Only
904 significant correlations are shown. Adjusted *P*-value indicates significance correction for
905 multiple treatment comparisons according to (Benjamini & Hochberg 1995). Significance, *** =
906 0.001, ** = 0.01, * = 0.05.

907 **Table S8**. Correlation results (Kendall's *tau*) for intracellular metabolites to *Calanus* grazing
908 ratios calculated from AmpliconNoise OTUs grouped into lower-level taxonomic ranks.
909 Metabolite - feature identification number; Correlation - Kendall's *tau* correlation test statistic;
910 Adj. P-value - significance adjusted for multiple treatment comparisons. Question marks
911 preceding metabolite identifications indicate degree of uncertainty, and represent the reverse
912 match factor of each metabolite when searched against the National Institute of Standards and
913 Technology chemistry database: from 750 to 850 (?), 650 to 750 (??), or 550 to 650 (???).

914 **Figure 1**. The 20 most abundant taxa present in *Calanus* gut content ("Copepod") and mesocosm
915 seawater ("Water") samples and their relative abundance in the metabarcoding data (excluding
916 *Calanus* sequences). (A) AmpliconNoise OTUs, (B) CREST fifth-rank taxonomic assignments.
917 Taxonomic assignments of OTUs and of sequence reads were performed using the CREST
918 classifier with SilvaMod reference database.

919 **Figure 2**. Non-metric multidimensional scaling (NMDS) ordination of (A) AmpliconNoise OTU
920 diversity and (B) CREST taxonomic diversity. *Calanus*-like reads were removed from the

921 sequence data prior to ordination. Dashed ellipses indicate 95% confidence intervals.

922 **Figure 3**. Comparison of methods used for quantification of (A) *Phaeocystis pouchetii* and (B)
923 *Skeletonema marinoi* in mesocosm water. Methods used were direct taxonomic classification of
924 pyrosequencing amplicons (Metabarcoding), cell counts per ml (Microscopy), and qPCR
925 (qPCR). Units shown are reads per ml, cells per ml, or SSU rRNA gene copies per ml,
926 respectively, for Metabarcoding, Microscopy and qPCR. Note logarithmic *y*-axis. Microscopy
927 and qPCR counts are from (Ray *et al*. 2016).

928 **Figure 4**. Numerical distribution of (A) AmpliconNoise OTU or (B) CREST taxonomic
929 assignment non-zero, finite grazing ratios (copepod gut content / seawater). The distribution of
930 grazing ratio values for all treatments and Raunefjorden into six arbitrary numerical categories,
931 indicated in the legend at the right, indicate various degrees of putatively "positive" (>1) or
932 "negative" (<1) grazing selection by *Calanus*. Full taxonomic content of AmpliconNoise OTU
933 and CREST grazing ratio tables can be found in the Supporting Information (Table S4).

934 **Figure 5**. Heatmap of metabolites, which are significant (correlation R > 0.235 of CAPdiscrim
935 of the peaksum normalized data set, Pearson's correlation test used for determination of
936 significance level) for the separation ("early", "middle", "late") in both NP and NPSi treatments.
937 Black=highest concentration, white=lowest concentration, Met=number of metabolite,
938 Ion=iontrace used for quantification, RT=retention time, tag "?", "??", or "???"=reverse match
939 factor of NIST database between 750 and 850, 650 and 750, or 550 and 650.compound classes:
940 CH=carbohydrates and derivatives, LP=lipides, TP=terpenoides, ST=steroles.

941 **Figure 6**. Clustering of intracellular metabolite profiles according to mesocosm succession phase
942 using canonical analysis of principle components using linear discriminant analysis

(CAPdiscrim). (A) NP mesocosms, Eigenvalues=0.95 and 0.40, Mis-classification error=25.7 %, Permutation test: $P = 0.001$ with 1000 permutations (B) NPSi mesocosms, Eigenvalues=0.96 and 0.40, Mis-classification error=22.5 %, Permutation test: $P = 0.001$ with 1000 permutations.

Figure S1. Contribution of copepod and putative symbiont reads to sequence datasets. Height of bars represents the total number of filtered 454 reads used for OTU clustering or taxonomic identification. The number of copepod OTUs/reads in each sample are indicated by black bars, while putative symbion OTUs/reads are shown as grey bars. Percentage of reads remaining after exclusion of copepod and putative symbiont OTUs/reads is shown above each column. (A) AmpliconNoise OTUs, (B) CREST taxonomic classifications.

Figure S2. Rarefaction analysis of 454 sequence data from *Calanus* gut content ("COP", A and C) and seawater ("FIL", B and D) samples. OTU clustering (A and B) was performed with a 98% similarity cut-off using AmpliconNoise v.1.29 with otherwise default parameters. Taxonomic assignments were performed using the CREST classifier and the SilvaMod database as reference taxonomy (C and D). Rarefaction is based on non-normalized read counts and a sampling frequency of 100 (A and B) or 10 (C and D). Treatment labels: blank, Control mesocosm; NP, NP mesocosm; NPSi, NPSi mesocosm; fjord, Raunefjorden. Numbers in sample labels indicate date of sampling in March 2012.

Figure S3. Microbial eukaryotes that distinguish *Calanus* gut content diversity from mesocosm seawater diversity when abundance normalized fifth-rank CREST taxonomic assignments were analyzed using Kruskal-Wallis rank sum tests (alpha = 0.001) with Sample Type (copepod or seawater) as grouping variable and signficance correction for multiple comparisons (q-value) (Benjamini & Hochberg 2995). Copepod sequences were removed prior to analysis.

965 **Figure S4**. Dynamics of *Calanus* sp. feeding selection over time in three mesocosm treatments
966 (Control, NP and NPSi) and Raunefjorden as assessed by metabarcoding analysis. The height of
967 bars (vertically centered at grazing ratio = 1, dashed black line) indicates ratios of the relative
968 abundance of taxonomic groups in copepod gut content divided by their relative abundance in
969 co-occurring plankton communities. Grazing ratios > 1 indicate higher abundance of a taxon in
970 *Calanus* gut content relative to taxon abundance in co-occuring plankton communities, while
971 grazing ratios < 1 indicate lower abundance in gut content relative to co-occurring plankton
972 communities. Sampling date (March 2012) is shown on the *x*-axis. (A) AmpliconNoise OTUs
973 grouped by fifth-rank taxonomic classification, (B) Fifth-rank taxonomic classifications.

Table 1.

Name	Sequence (5'-3')[1]	T_{an2}	Position[3]	Reference
F-1183mod	AATTTGACTCAACRCGGG	60.2	1183-1200	Hadziavdic *et al.* 2014
R-1443mod	GRGCATCACAGACCTG		1443-1428	Hadziavdic *et al.* 2014
Cal-SpcC3-block	CTGTTATTGCTCAATCTY GTGCGAC[SpcC3]	70.2	1430-1406	This study
Cal-PNA-block	[NH_2]-CTAAGAGTCGCCA GTCCC-[COOH]	70.2	1406-1389	This study

[1] [SpcC3], -C-C-C-OH; [NH_2], N-terminus of peptide backbone; [COOH], C-terminus of peptide backbone

[2] Annealing temperature used for PCR, in degrees Celsius

[3] Numbering based on reference alignment described in Hadziavdic *et al.* 2014

Table 2.

Pipeline	Distance metric	Explanatory variable	F-value	P-value	R^2
Amplicon Noise	Jaccard	Sample Type	13.279	**0.001**	0.210
		Experiment	2.420	**0.017**	0.038
		Treatment	0.549	0.996	0.017
		Date	4.385	**0.001**	0.069
	Bray-Curtis	Sample Type	22.592	**0.001**	0.300
		Experiment	3.454	**0.003**	0.046
		Treatment	0.443	0.996	0.012
		Date	6.279	**0.001**	0.084
CREST taxonomic assignment	Jaccard	Sample Type	30.930	**0.001**	0.362
		Experiment	1.170	0.267	0.014
		Treatment	1.754	0.080	0.041
		Date	7.919	**0.001**	0.093
	Bray-Curtis	Sample Type	51.043	**0.001**	0.470
		Experiment	1.151	0.286	0.011
		Treatment	1.986	0.071	0.037
		Date	10.465	**0.002**	0.096

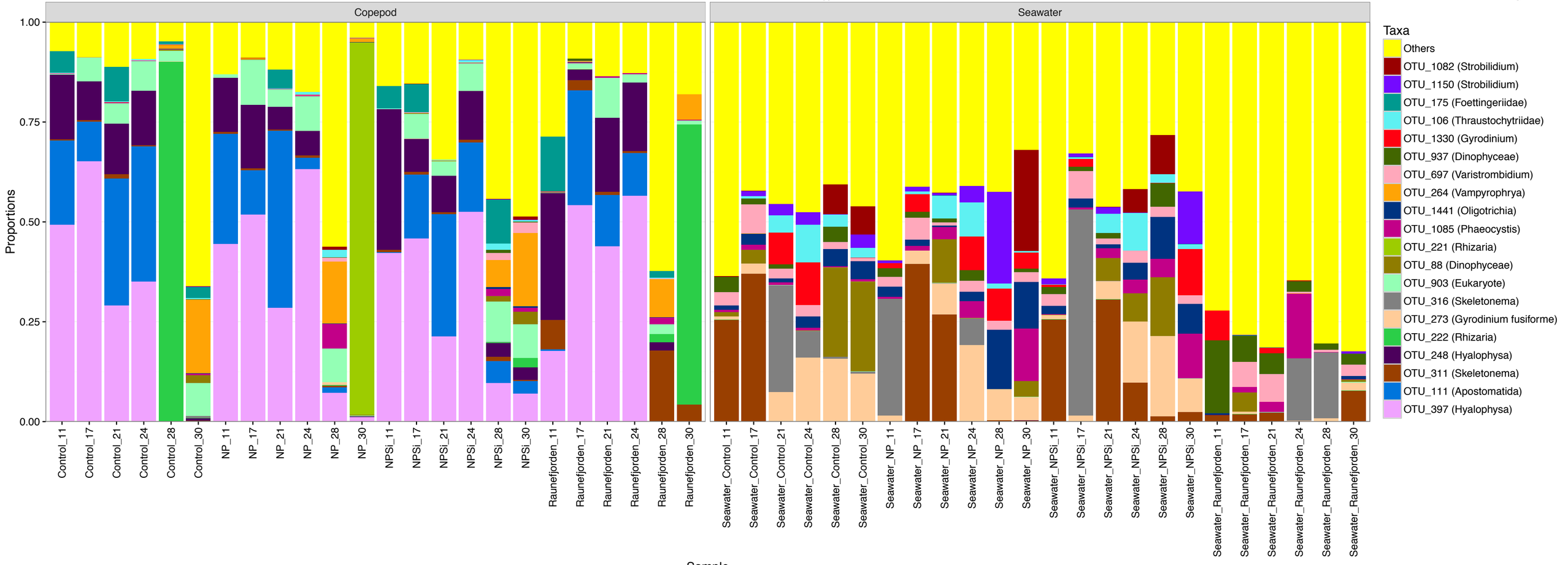

Copepod
Seawater
Proportions
1.00
0.75
0.50
0.25
0.00
Control_11
Control_17
Control_21
Control_24
Control_28
Control_30
NP_11
NP_17
NP_21
NP_24
NP_28
NP_30
NPSi_11
NPSi_17
NPSi_21
NPSi_24
NPSi_28
NPSi_30
Raunefjorden_11
Raunefjorden_17
Raunefjorden_21
Raunefjorden_24
Raunefjorden_28
Raunefjorden_30
Seawater_Control_11
Seawater_Control_17
Seawater_Control_21
Seawater_Control_24
Seawater_Control_28
Seawater_Control_30
Seawater_NP_11
Seawater_NP_17
Seawater_NP_21
Seawater_NP_24
Seawater_NP_28
Seawater_NP_30
Seawater_NPSi_11
Seawater_NPSi_17
Seawater_NPSi_21
Seawater_NPSi_24
Seawater_NPSi_28
Seawater_NPSi_30
Seawater_Raunefjorden_11
Seawater_Raunefjorden_17
Seawater_Raunefjorden_21
Seawater_Raunefjorden_24
Seawater_Raunefjorden_28
Seawater_Raunefjorden_30
Sample
Taxa
Others
OTU_1082 (Strobilidium)
OTU_1150 (Strobilidium)
OTU_175 (Foettingeriidae)
OTU_106 (Thraustochytriidae)
OTU_1330 (Gyrodinium)
OTU_937 (Dinophyceae)
OTU_697 (Varistrombidium)
OTU_264 (Vampyrophrya)
OTU_1441 (Oligotrichia)
OTU_1085 (Phaeocystis)
OTU_221 (Rhizaria)
OTU_88 (Dinophyceae)
OTU_903 (Eukaryote)
OTU_316 (Skeletonema)
OTU_273 (Gyrodinium fusiforme)
OTU_222 (Rhizaria)
OTU_248 (Hyalophysa)
OTU_311 (Skeletonema)
OTU_111 (Apostomatida)
OTU_397 (Hyalophysa)

Figure 1B

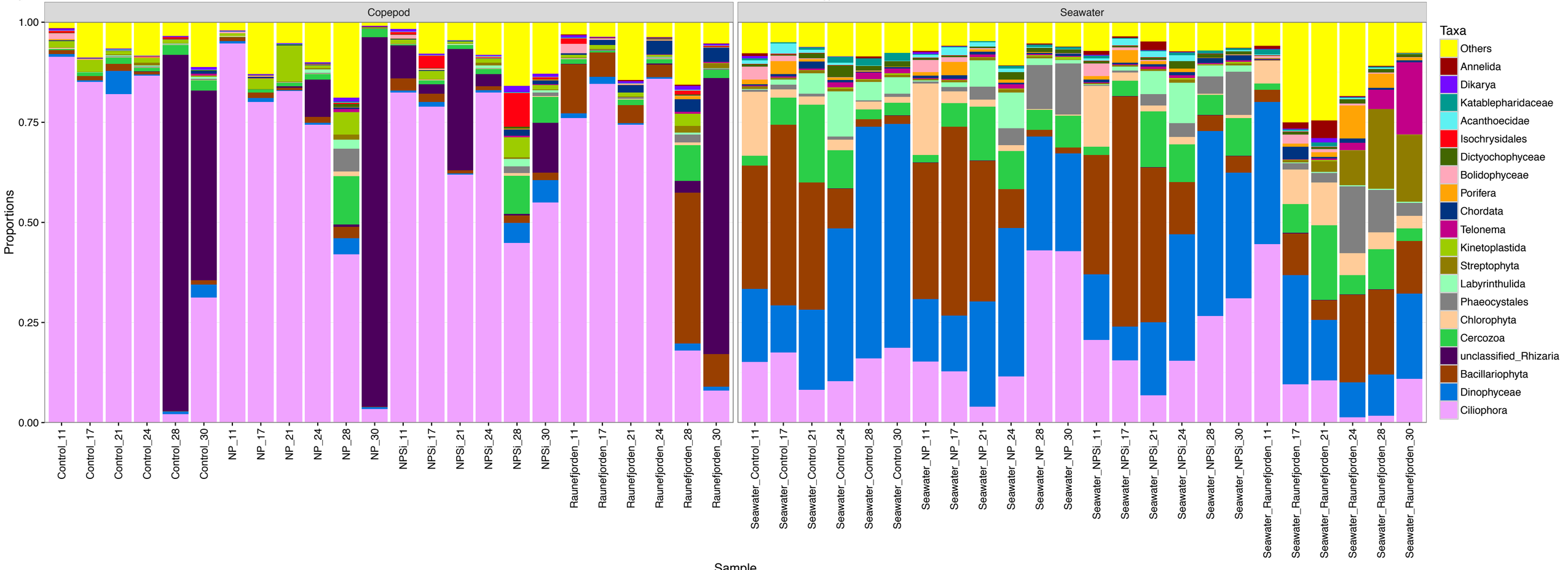

Copepod
Seawater
Proportions
1.00
0.75
0.50
0.25
0.00
Sample
Taxa
Others
Annelida
Dikarya
Katablepharidaceae
Acanthoecidae
Isochrysidales
Dictyochophyceae
Bolidophyceae
Porifera
Chordata
Telonema
Kinetoplastida
Streptophyta
Labyrinthulida
Phaeocystales
Chlorophyta
Cercozoa
unclassified_Rhizaria
Bacillariophyta
Dinophyceae
Ciliophora
Control_11
Control_17
Control_21
Control_24
Control_28
Control_30
NP_11
NP_17
NP_21
NP_24
NP_28
NP_30
NPSi_11
NPSi_17
NPSi_21
NPSi_24
NPSi_28
NPSi_30
Raunefjorden_11
Raunefjorden_17
Raunefjorden_21
Raunefjorden_24
Raunefjorden_28
Raunefjorden_30
Seawater_Control_11
Seawater_Control_17
Seawater_Control_21
Seawater_Control_24
Seawater_Control_28
Seawater_Control_30
Seawater_NP_11
Seawater_NP_17
Seawater_NP_21
Seawater_NP_24
Seawater_NP_28
Seawater_NP_30
Seawater_NPSi_11
Seawater_NPSi_17
Seawater_NPSi_21
Seawater_NPSi_24
Seawater_NPSi_28
Seawater_NPSi_30
Seawater_Raunefjorden_11
Seawater_Raunefjorden_17
Seawater_Raunefjorden_21
Seawater_Raunefjorden_24
Seawater_Raunefjorden_28
Seawater_Raunefjorden_30

Figure 2A

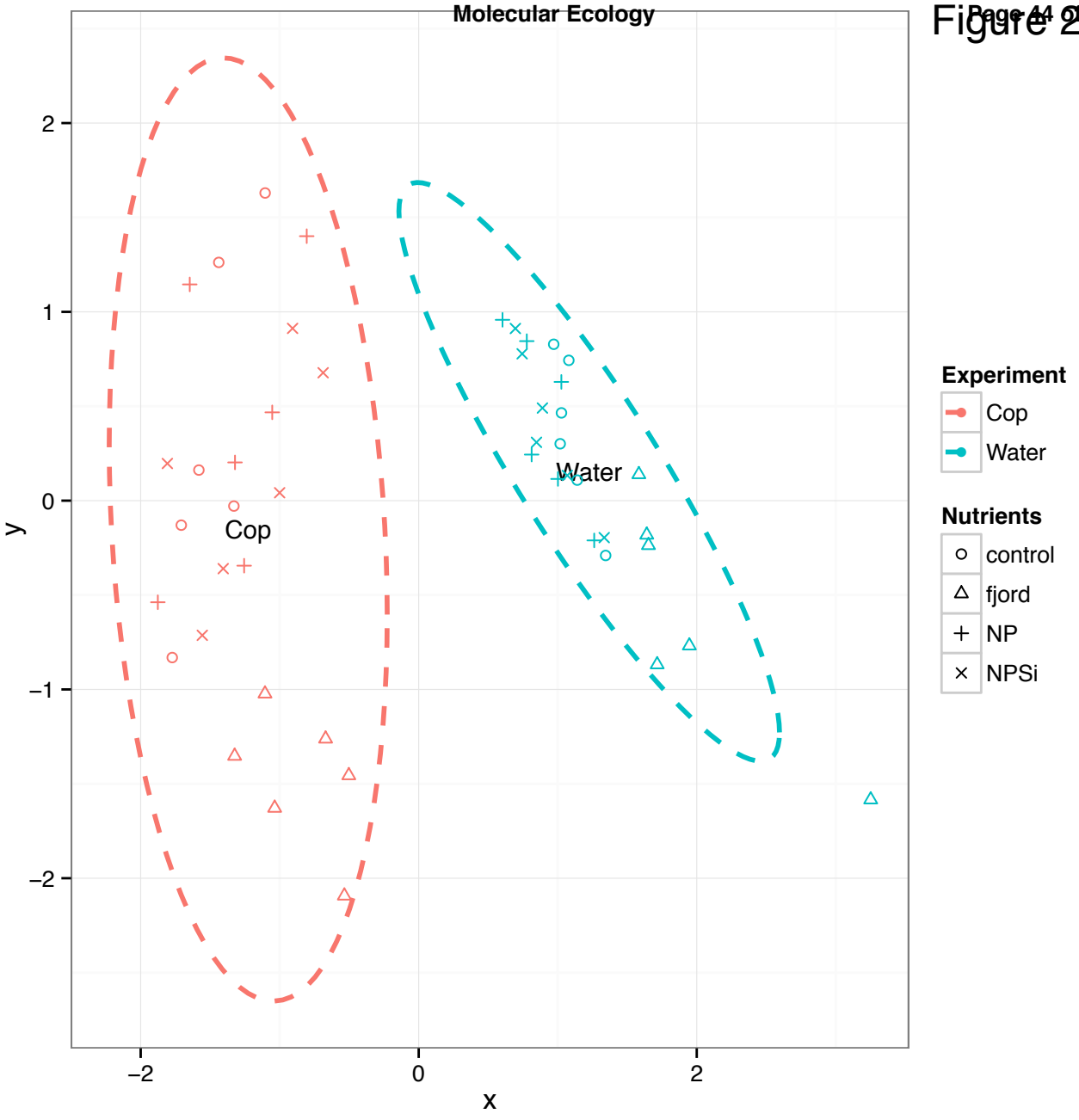

Molecular Ecology
Cop
Water
y
x
2
1
0
−1
−2
−2
0
2
Experiment
Cop
Water
Nutrients
control
fjord
NP
NPSi

Figure 2B

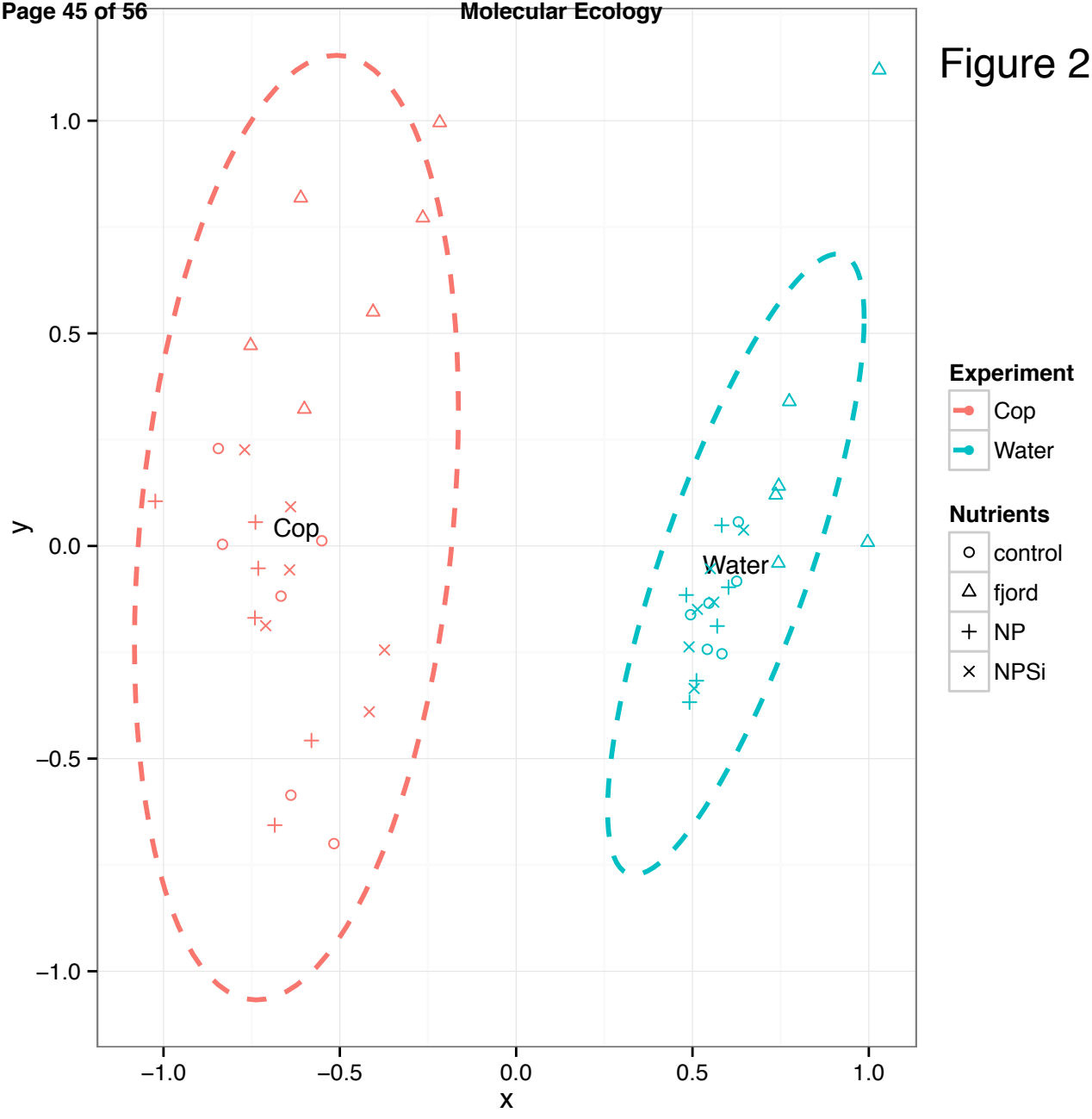

Cop
Water
y
x
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
Experiment
Cop
Water
Nutrients
control
fjord
NP
NPSi

Figure 3

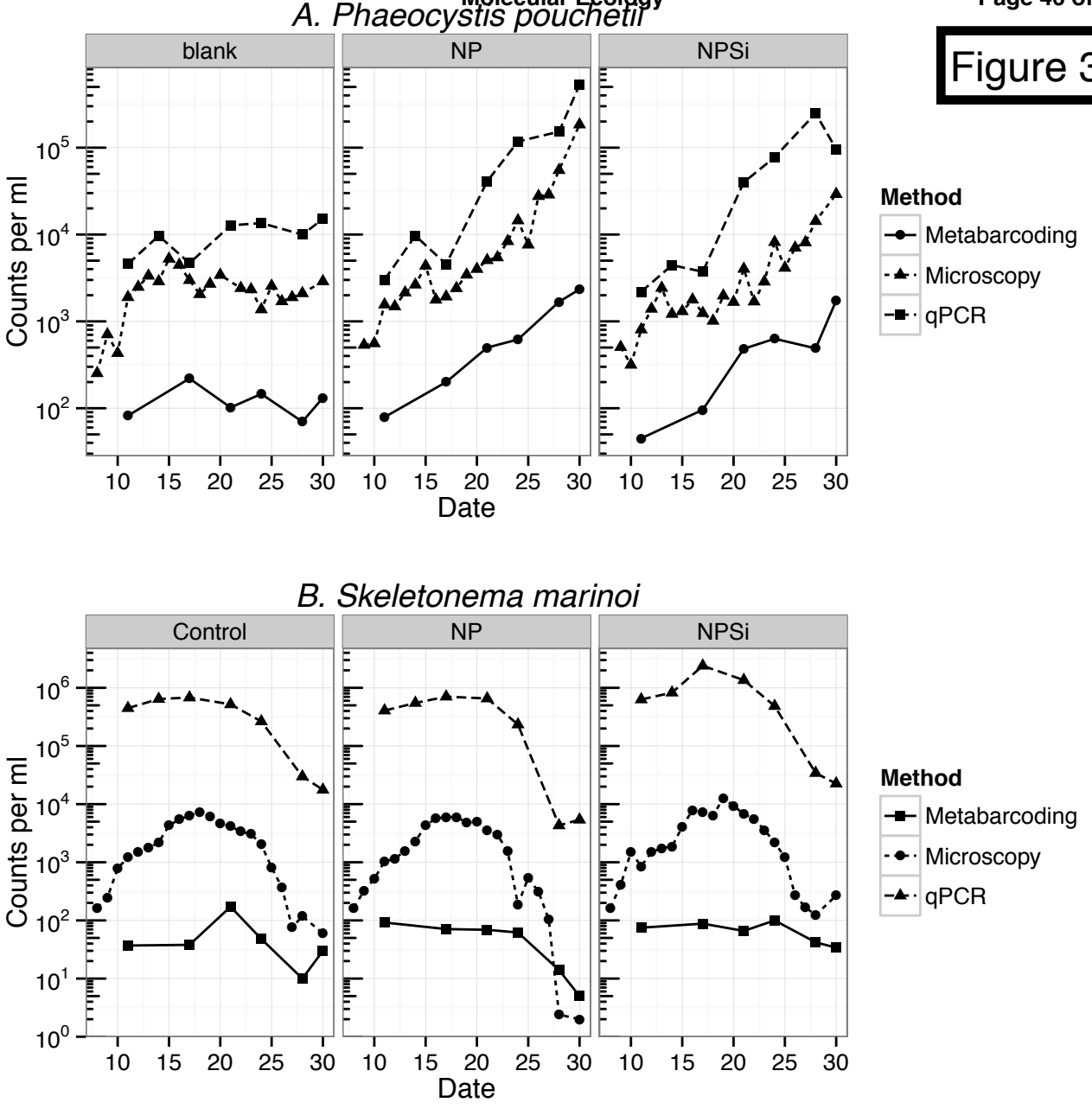

A. Phaeocystis pouchetii
blank
NP
NPSi
Counts per ml
Date
Method
Metabarcoding
Microscopy
qPCR
B. Skeletonema marinoi
Control
NP
NPSi
Counts per ml
Date
Method
Metabarcoding
Microscopy
qPCR

Figure 4A. AmpliconNoise OTU ratios

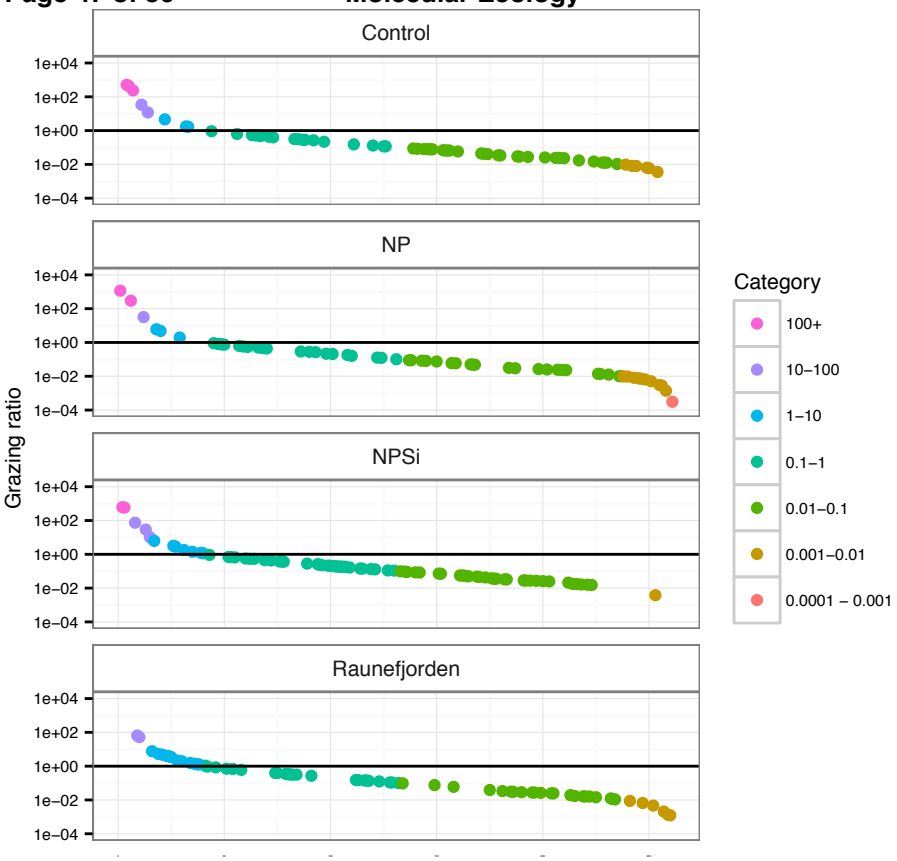


Figure 4B. CREST taxonomy ratios

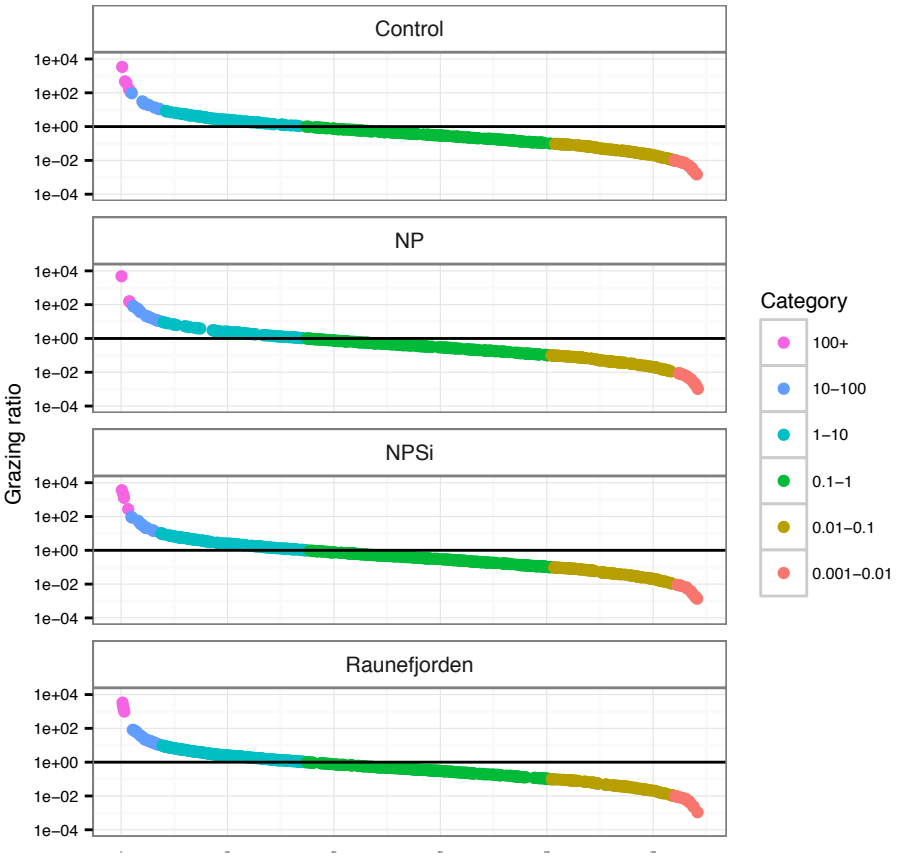
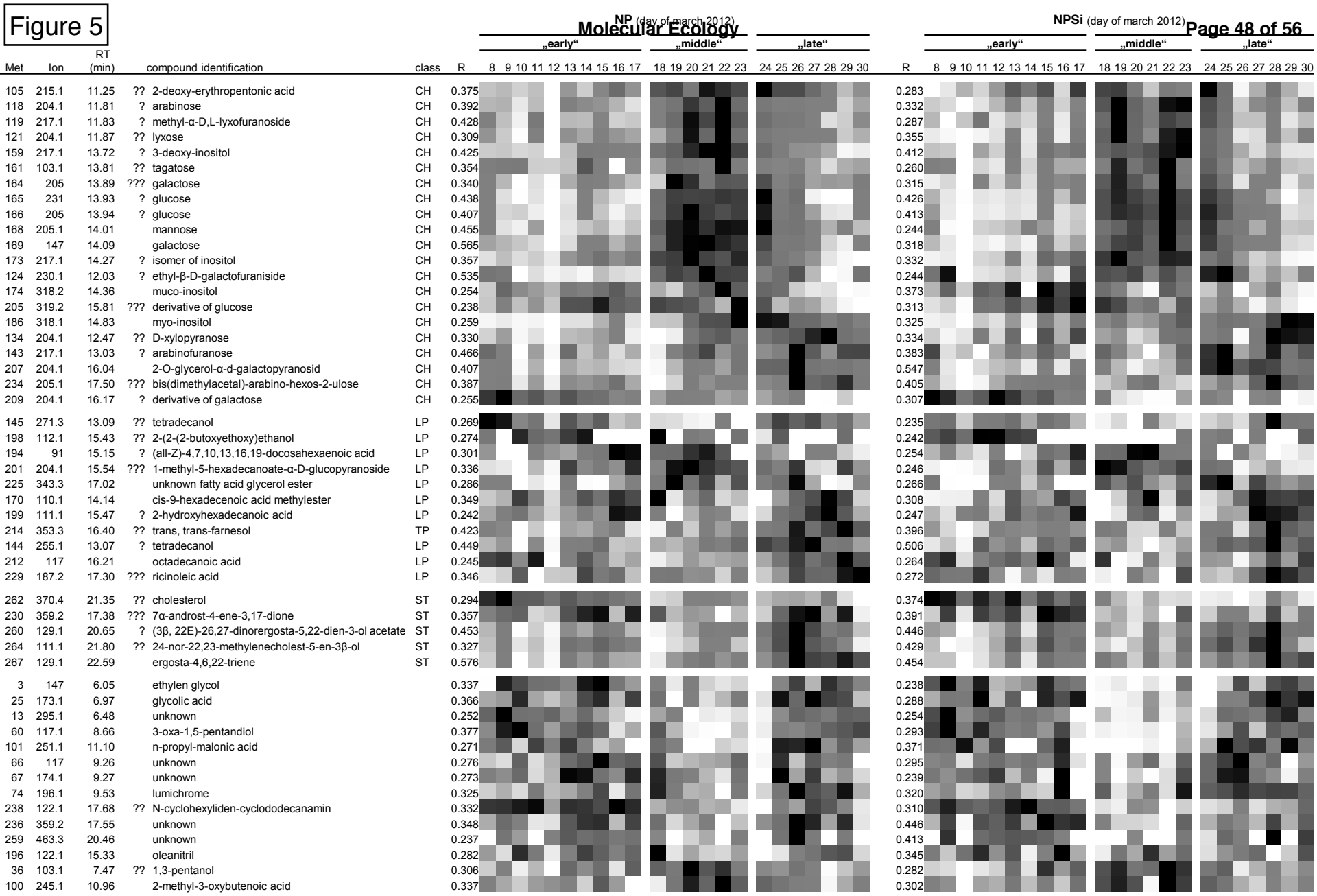


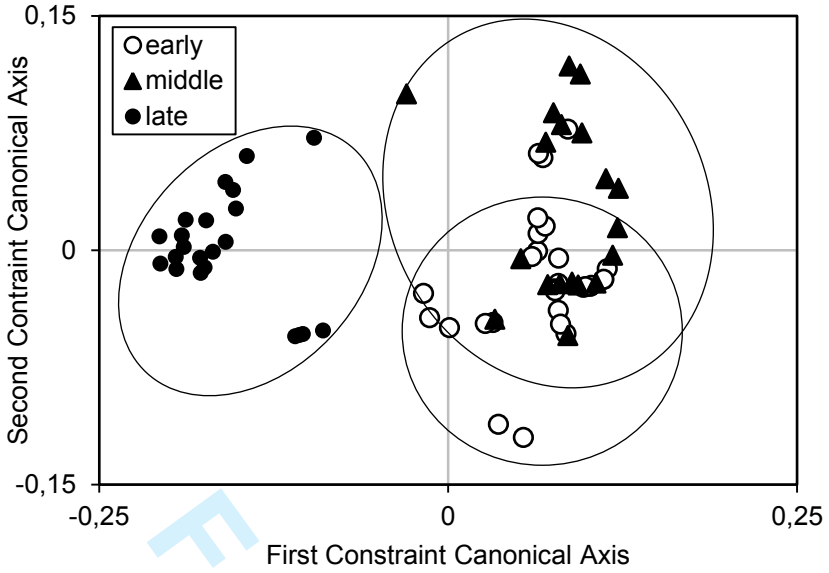
Figure 5



Met	Ion	RT (min)		compound identification	class	R	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	R	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
105	215.1	11.25	??	2-deoxy-erythropentonic acid	CH	0.375																								0.283																							
118	204.1	11.81	?	arabinose	CH	0.392																								0.332																							
119	217.1	11.83	?	methyl-α-D,L-lyxofuranoside	CH	0.428																								0.287																							
121	204.1	11.87	??	lyxose	CH	0.309																								0.355																							
159	217.1	13.72	?	3-deoxy-inositol	CH	0.425																								0.412																							
161	103.1	13.81	??	tagatose	CH	0.354																								0.260																							
164	205	13.89	???	galactose	CH	0.340																								0.315																							
165	231	13.93	?	glucose	CH	0.438																								0.426																							
166	205	13.94	?	glucose	CH	0.407																								0.413																							
168	205.1	14.01		mannose	CH	0.455																								0.244																							
169	147	14.09		galactose	CH	0.565																								0.318																							
173	217.1	14.27	?	isomer of inositol	CH	0.357																								0.332																							
124	230.1	12.03	?	ethyl-β-D-galactofuraniside	CH	0.535																								0.244																							
174	318.2	14.36		muco-inositol	CH	0.254																								0.373																							
205	319.2	15.81	???	derivative of glucose	CH	0.238																								0.313																							
186	318.1	14.83		myo-inositol	CH	0.259																								0.325																							
134	204.1	12.47	??	D-xylopyranose	CH	0.330																								0.334																							
143	217.1	13.03	?	arabinofuranose	CH	0.466																								0.383																							
207	204.1	16.04		2-O-glycerol-α-d-galactopyranosid	CH	0.407																								0.547																							
234	205.1	17.50	???	bis(dimethylacetal)-arabino-hexos-2-ulose	CH	0.387																								0.405																							
209	204.1	16.17	?	derivative of galactose	CH	0.255																								0.307																							
145	271.3	13.09	??	tetradecanol	LP	0.269																								0.235																							
198	112.1	15.43	??	2-(2-(2-butoxyethoxy)ethanol	LP	0.274																								0.242																							
194	91	15.15	?	(all-Z)-4,7,10,13,16,19-docosahexaenoic acid	LP	0.301																								0.254																							
201	204.1	15.54	???	1-methyl-5-hexadecanoate-α-D-glucopyranoside	LP	0.336																								0.246																							
225	343.3	17.02		unknown fatty acid glycerol ester	LP	0.286																								0.266																							
170	110.1	14.14		cis-9-hexadecenoic acid methylester	LP	0.349																								0.308																							
199	111.1	15.47	?	2-hydroxyhexadecanoic acid	LP	0.242																								0.247																							
214	353.3	16.40	??	trans, trans-farnesol	TP	0.423																								0.396																							
144	255.1	13.07	?	tetradecanol	LP	0.449																								0.506																							
212	117	16.21		octadecanoic acid	LP	0.245																								0.264																							
229	187.2	17.30	???	ricinoleic acid	LP	0.346																								0.272																							
262	370.4	21.35	??	cholesterol	ST	0.294																								0.374																							
230	359.2	17.38	???	7α-androst-4-ene-3,17-dione	ST	0.357																								0.391																							
260	129.1	20.65	?	(3β, 22E)-26,27-dinorergosta-5,22-dien-3-ol acetate	ST	0.453																								0.446																							
264	111.1	21.80	??	24-nor-22,23-methylenecholest-5-en-3β-ol	ST	0.327																								0.429																							
267	129.1	22.59		ergosta-4,6,22-triene	ST	0.576																								0.454																							
3	147	6.05		ethylen glycol		0.337																								0.238																							
25	173.1	6.97		glycolic acid		0.366																								0.288																							
13	295.1	6.48		unknown		0.252																								0.254																							
60	117.1	8.66		3-oxa-1,5-pentandiol		0.377																								0.293																							
101	251.1	11.10		n-propyl-malonic acid		0.271																								0.371																							
66	117	9.26		unknown		0.276																								0.295																							
67	174.1	9.27		unknown		0.273																								0.239																							
74	196.1	9.53		lumichrome		0.325																								0.320																							
238	122.1	17.68	??	N-cyclohexyliden-cyclododecanamin		0.332																								0.310																							
236	359.2	17.55		unknown		0.348																								0.446																							
259	463.3	20.46		unknown		0.237																								0.413																							
196	122.1	15.33		oleanitril		0.282																								0.345																							
36	103.1	7.47	??	1,3-pentanol		0.306																								0.282																							
100	245.1	10.96		2-methyl-3-oxybutenoic acid		0.337																								0.302																							

Figure 6

A



B

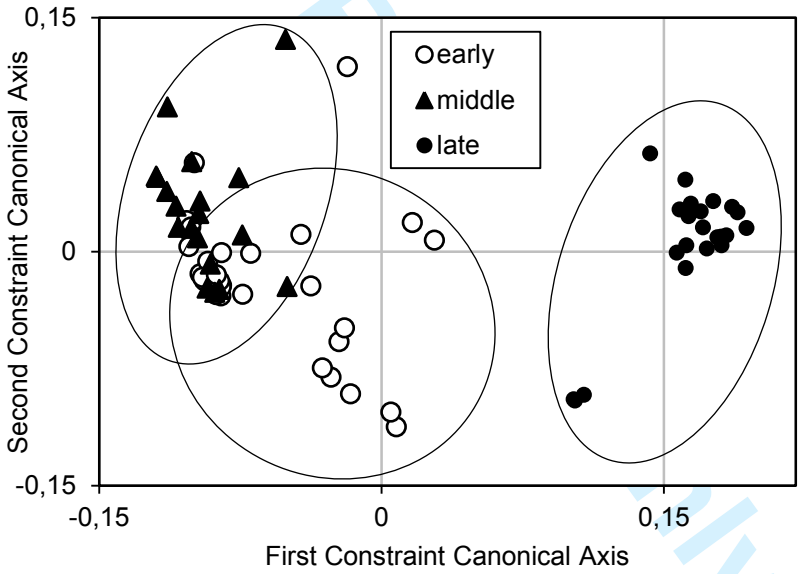


Figure S1A

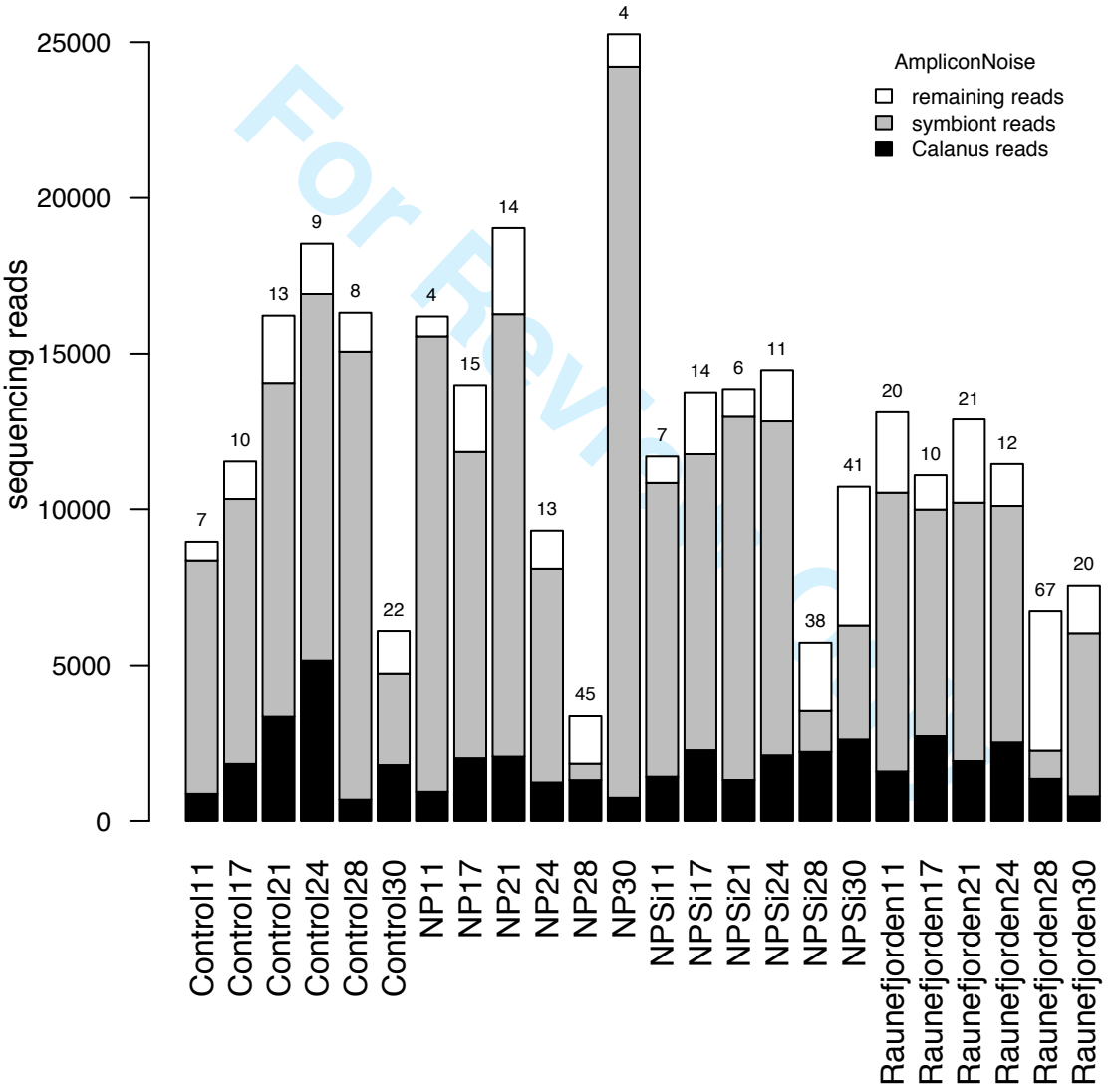
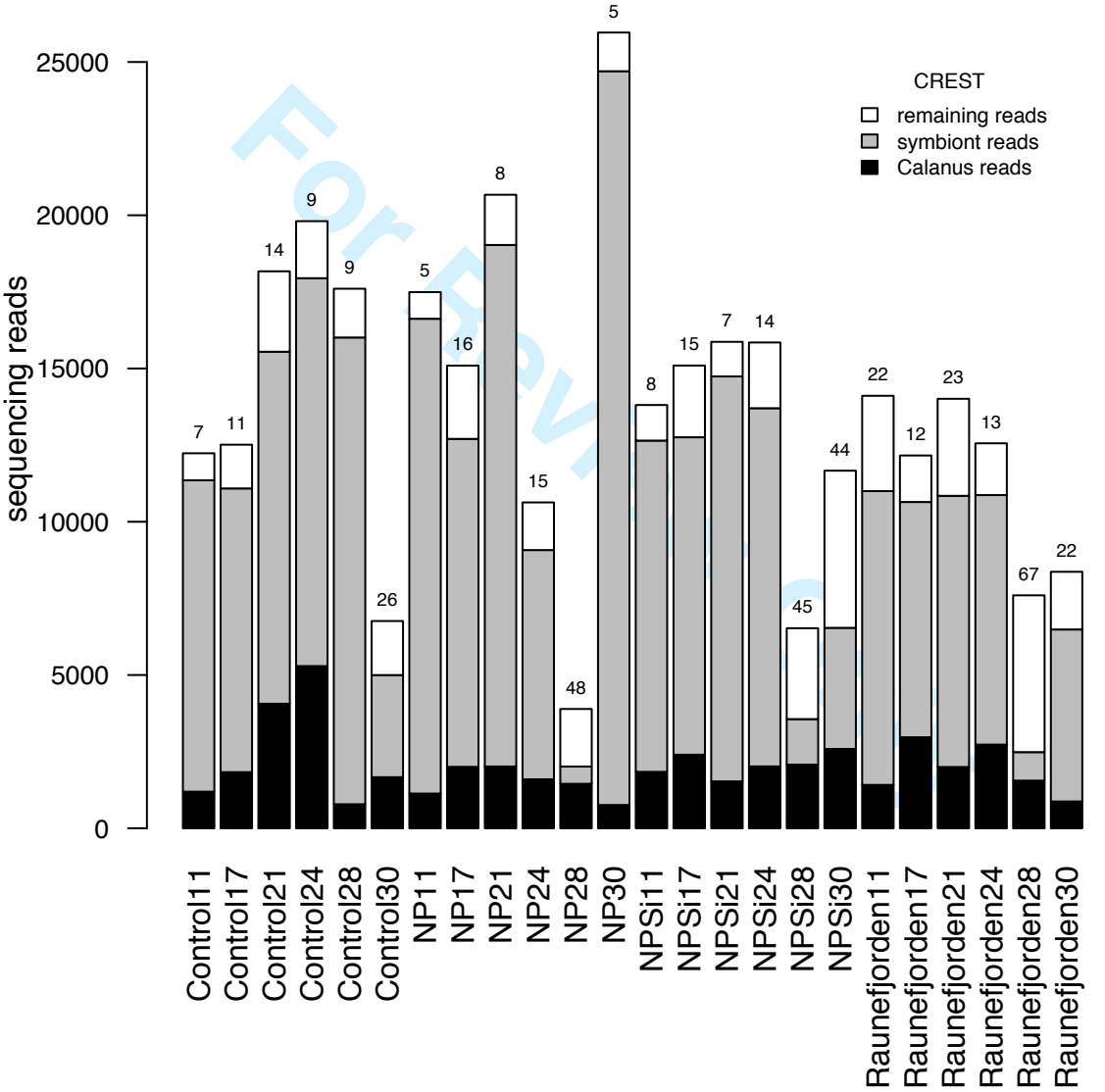
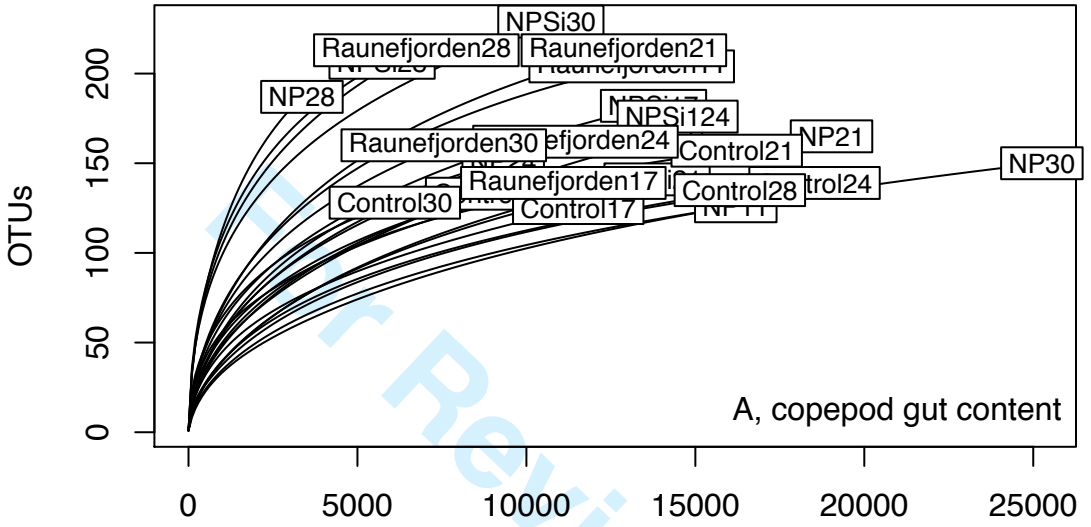


Figure S1B





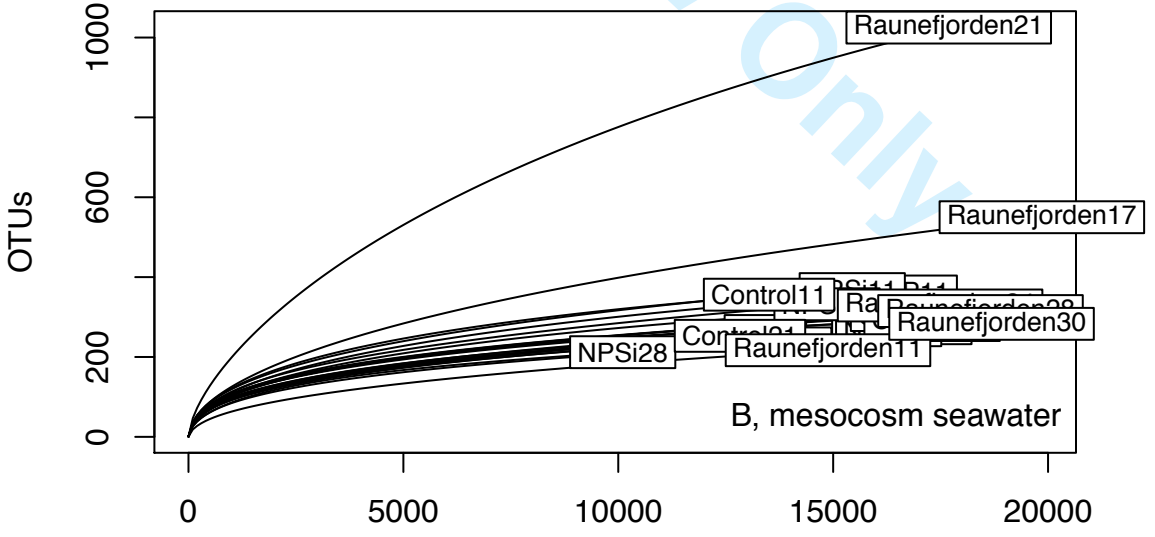
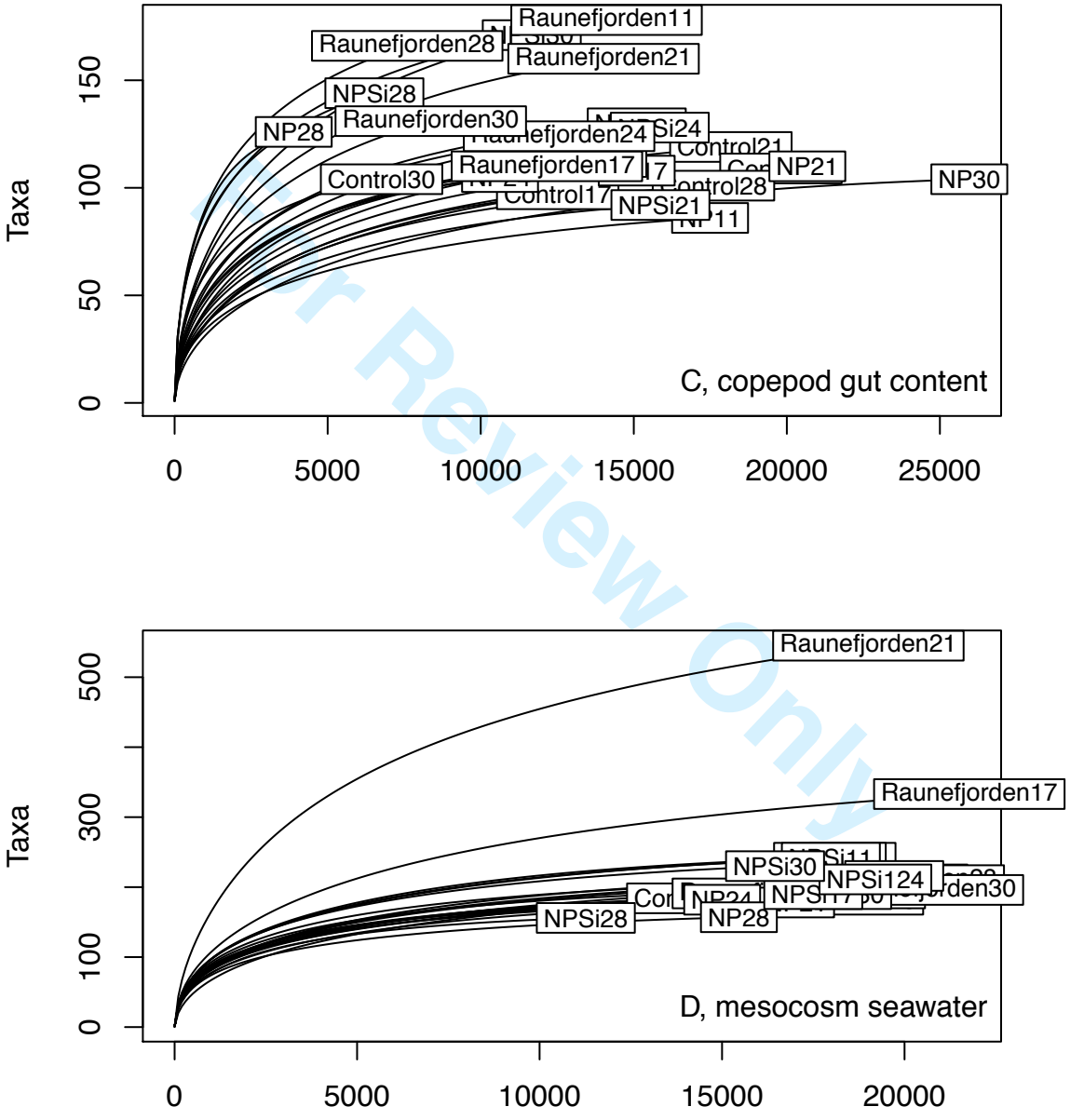


Figure S2C and S2D



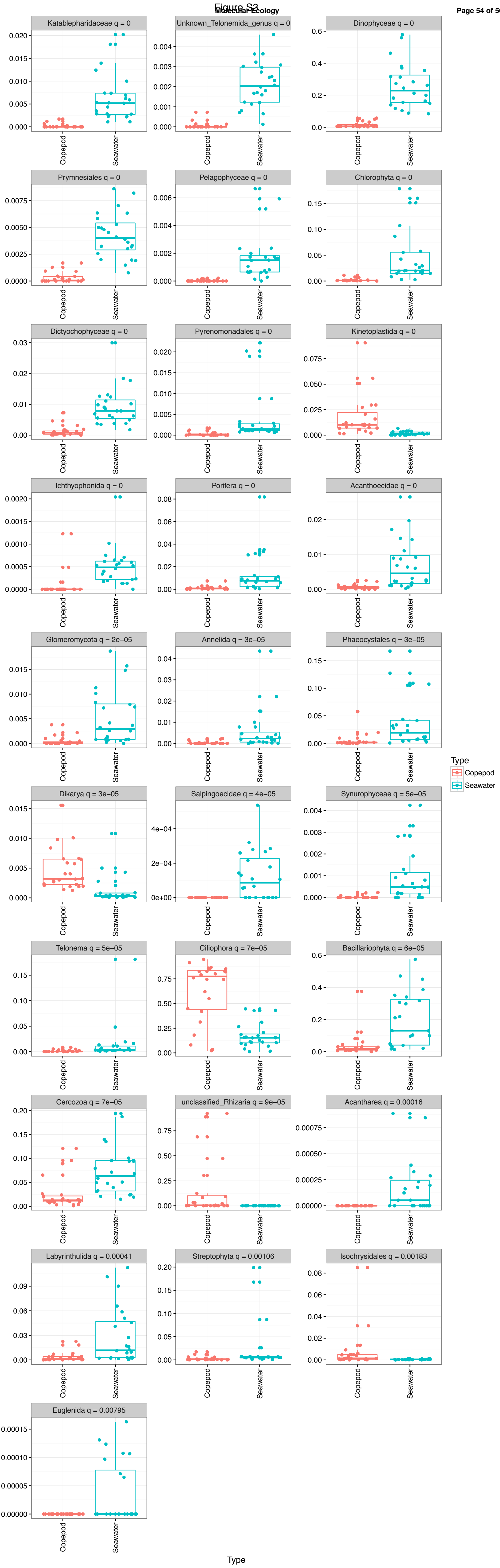
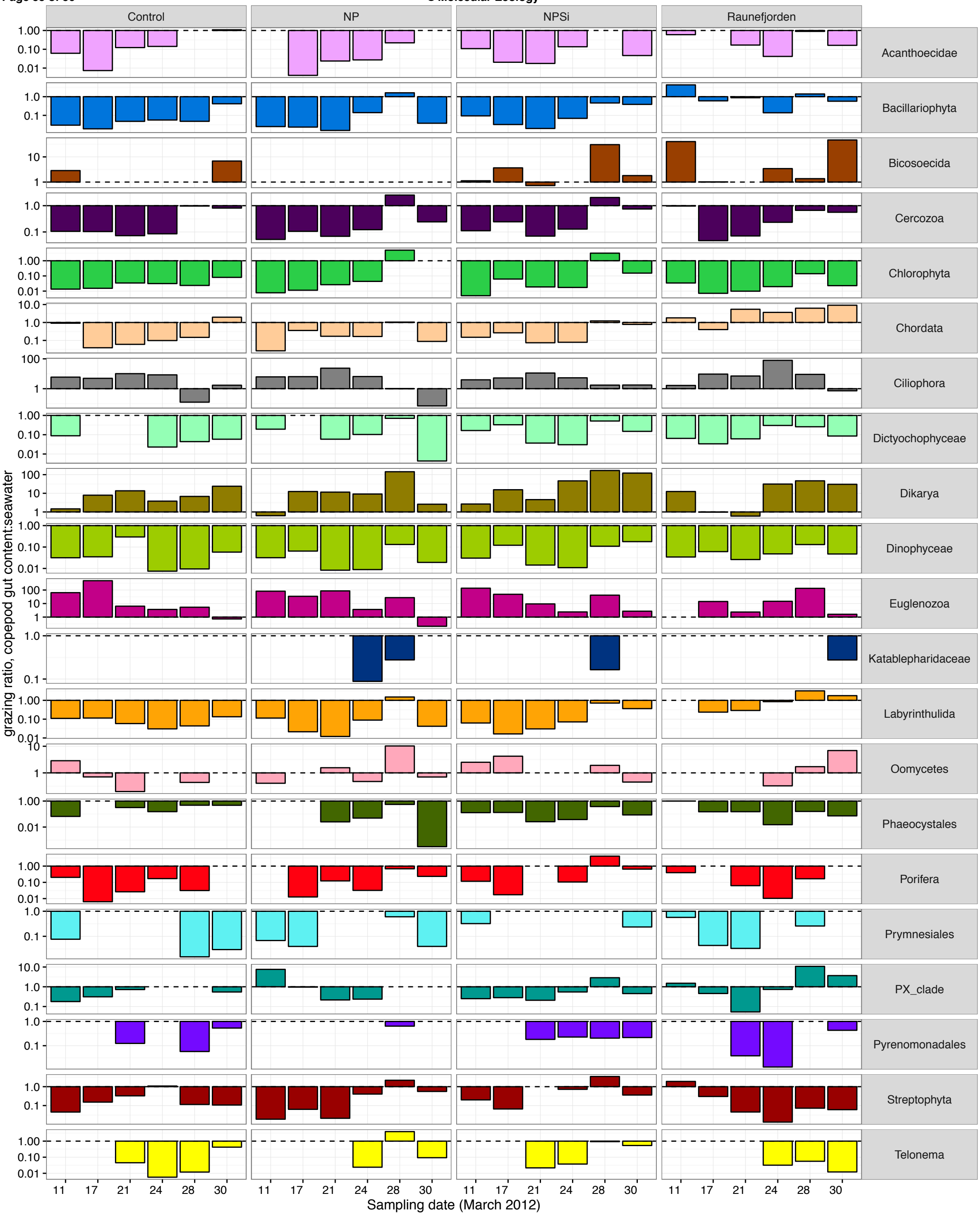
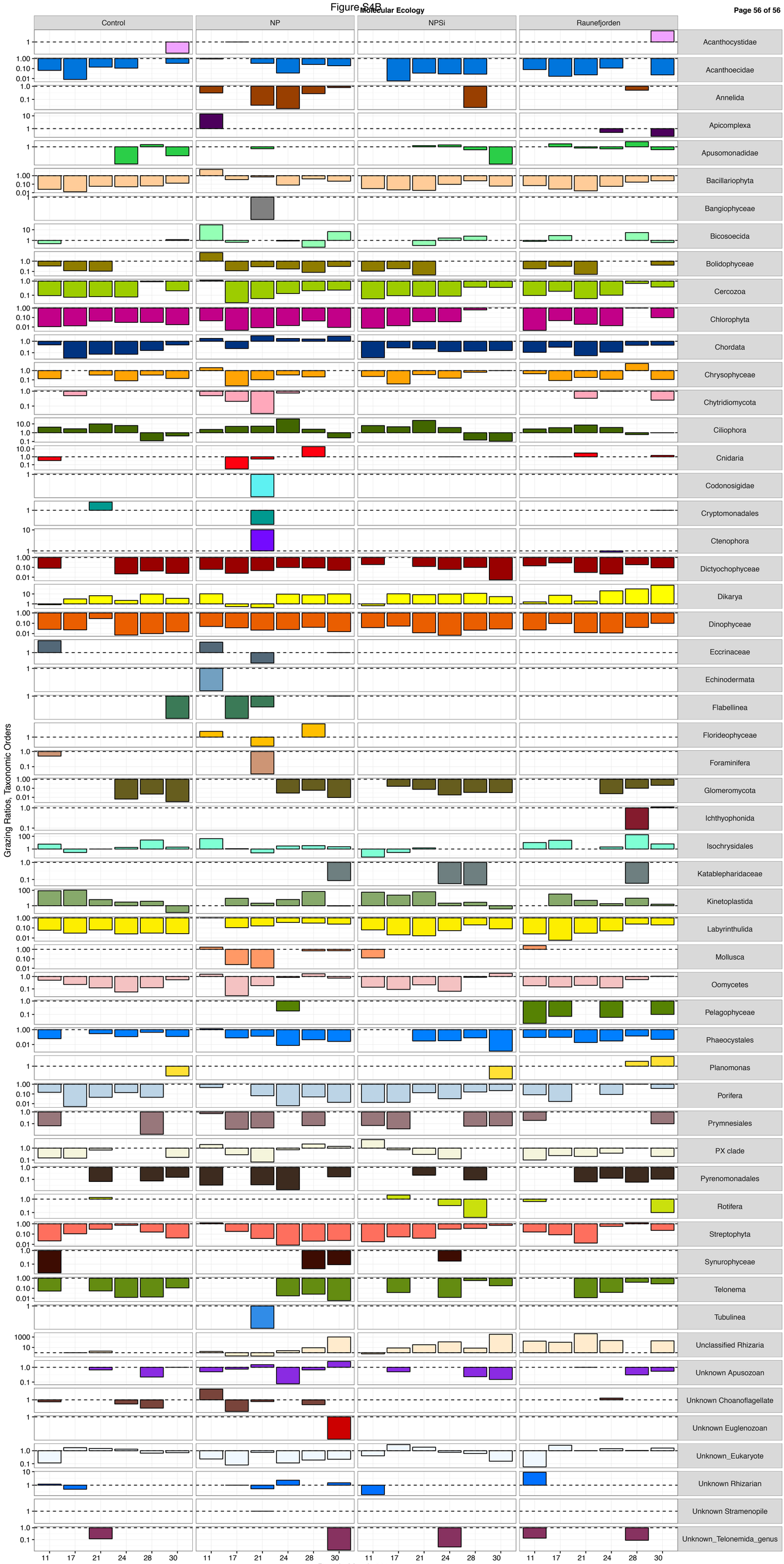

Katablepharidaceae q = 0
Unknown_Telonemida_genus q = 0
Dinophyceae q = 0
Prymnesiales q = 0
Pelagophyceae q = 0
Chlorophyta q = 0
Dictyochophyceae q = 0
Pyrenomonadales q = 0
Kinetoplastida q = 0
Ichthyophonida q = 0
Porifera q = 0
Acanthoecidae q = 0
Glomeromycota q = 2e−05
Annelida q = 3e−05
Phaeocystales q = 3e−05
Dikarya q = 3e−05
Salpingoecidae q = 4e−05
Synurophyceae q = 5e−05
Telonema q = 5e−05
Ciliophora q = 7e−05
Bacillariophyta q = 6e−05
Cercozoa q = 7e−05
unclassified_Rhizaria q = 9e−05
Acantharea q = 0.00016
Labyrinthulida q = 0.00041
Streptophyta q = 0.00106
Isochrysidales q = 0.00183
Euglenida q = 0.00795
Copepod
Seawater
Abundance normalized
Type
Copepod
Seawater
Type

Figure S4A


Control
NP
NPSi
Raunefjorden
Acanthoecidae
Bacillariophyta
Bicosoecida
Cercozoa
Chlorophyta
Chordata
Ciliophora
Dictyochophyceae
Dikarya
Dinophyceae
Euglenozoa
Katablepharidaceae
Labyrinthulida
Oomycetes
Phaeocystales
Porifera
Prymnesiales
PX_clade
Pyrenomonadales
Streptophyta
Telonema
grazing ratio, copepod gut content:seawater
11
17
21
24
28
30
Sampling date (March 2012)


Control
NP
NPSi
Raunefjorden
Acanthocystidae
Acanthoecidae
Annelida
Apicomplexa
Apusomonadidae
Bacillariophyta
Bangiophyceae
Bicosoecida
Bolidophyceae
Cercozoa
Chlorophyta
Chordata
Chrysophyceae
Chytridiomycota
Ciliophora
Cnidaria
Codonosigidae
Cryptomonadales
Ctenophora
Dictyochophyceae
Dikarya
Dinophyceae
Eccrinaceae
Echinodermata
Flabellinea
Florideophyceae
Foraminifera
Glomeromycota
Ichthyophonida
Isochrysidales
Katablepharidaceae
Kinetoplastida
Labyrinthulida
Mollusca
Oomycetes
Pelagophyceae
Phaeocystales
Planomonas
Porifera
Prymnesiales
PX clade
Pyrenomonadales
Rotifera
Streptophyta
Synurophyceae
Telonema
Tubulinea
Unclassified Rhizaria
Unknown Apusozoan
Unknown Choanoflagellate
Unknown Euglenozoan
Unknown_Eukaryote
Unknown Rhizarian
Unknown Stramenopile
Unknown_Telonemida_genus
Grazing Ratios, Taxonomic Orders
11 17 21 24 28 30
Date in March 2012